Energy Efficiency in Hybrid Mobile and Wireless Networks

Ziaul Haq Abbas

Energy Efficiency in Hybrid Mobile and Wireless Networks

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of *Philosophiae Doctor (PhD)* in Information and Communication Technology

Department of Information and Communication Technology
Faculty of Engineering and Science
University of Agder
2012

Doctoral Dissertation at the University of Agder 49
ISBN: 978-82-7117-712-6
ISSN: 1504-9272

Printed by the Printing Office, University of Agder
Kristiansand

To my beloved wife Ambreen, and our angels: Abia, Abdullah, and Ahmad

To my ever-sacrificing mother Waseem-un-Nissa, who taught me steadfastness

To the memories of my father Azizul Haq, who taught me patience

Preface

The research work reported in this dissertation was carried out at the Agder Mobility Laboratory (AML), Department of Information and Communication Technology (ICT), University of Agder (UiA), in Grimstad, Norway, from August 2008 to June 2012, including 30 credit points of the course study. My main supervisor has been Professor Frank Y. Li, University of Agder, and my co-supervisor has been Professor Andreas Prinz, University of Agder. From March 2012 to July 2012, I was a visiting researcher at the University of Minnesota (UMN), USA. My host professor at UMN was Professor Zhi-Quan (Tom) Luo.

The study was financed by the Higher Education Commission (HEC), Pakistan, under an overseas PhD scholarship scheme. My visit to UMN was financed by the European Commission under the 7th Framework Program (FP7) through the Security, Services, Networking, and Performance of Next Generation IP-Based Multimedia Wireless Networks Project (S2EuNet, http://s2eunet.org/) under Agreement Number 247083.

Production note: This dissertation as well as my papers produced during my PhD study have been written using LaTeX. The mathematical calculations and simulation results were obtained by using MATLAB.

Acknowledgements

First of all, I want to express my deepest gratitude to Almighty Allah, for His blessings. The strength and vision He bestowed upon me kept me focused throughout the course of this study. The successful completion of my PhD degree is just one of His endless blessings for which I am thoroughly obliged to.

I feel heavily indebted to my supervisor Prof. Frank Y. Li for teaching me the art of doing scientific work and the way of being a researcher. He always knew when to pull me out of the rare, though inevitable, pessimistic times. I cannot forget his appreciation of my academic as well as financial limitations, and his profound impact on my intellectual development. With his enthusiasm, diligence and inspiration he provided encouragement, sound advice and good teaching. I would also like to express my gratitude to my co-supervisor Prof. Andreas Prinz for his support and encouragement through these years. Prof. Frank Reichert also deserves my thanks for implicitly teaching me how to be persistent to achieve the goals. I am also grateful to Prof. Zhi-Quan (Tom) Luo for giving me the valuable opportunity to have a learning experience in USA. My graduate and undergraduate teachers, Prof. Ijaz Mansoor Qureshi and Nasir Jamil, respectively deserve my gratitude for providing me the impetus to pursue for this highest academic degree.

It is an honour for me to acknowledge my friends and colleagues at UiA, Ahmed Aboughonim, Batool Talha, Gulzaib Rafiq, Hongzhi Jiao, Jaziar Radianti, Lei Jiao, Terje Gjøsæter, and Xin He, for their support, company, and valuable discussions. The lively company and help from Ghislain Maurice will be remembered. I also appreciate the assistance provided by Tor Erik Christensen and Pål Grandal, from the International Office at UiA. The support for my PhD study provided by the Higher Education Commission (HEC) of Pakistan and the Norwegian Centre for International Cooperation in Education (SIU) is greatly appreciated.

This thesis gives me an immense sense of honour to mention the strength I got through my mother's prayers and great sacrifices for me. My father's polite attitude for the family is unforgettable. His words of inspiration and encouragement, in pursuit of excellence, are still with me and will last forever. I wish he were alive to witness the completion of my PhD. The encouragement and help from my sister Azra and brothers Zahoor, Ehsan, and Imran have given me impetus to move forward in my academic career. I would also like to thank Lala Majeed, Gulnaz Bhabhi, Tahira Bhabhi and Rabeea for their wishes and caring attitude.

I deeply acknowledge the encouragement, support, and prayers I had from Ghazala Baji, Nabeela (Mo) Baji, Sumbla, Naureen (Beena) and Waqar Bhai. Their sacrifices are very deeply appreciated. Without the major role played by these five persons, my family reunion in Norway with my wife and kids would have been an impossible task to accomplish. I owe them a lot of gratitude.

Last but not least, I thank my wife, Ambreen for her great patience, sacrifices, unending support and encouragement throughout the period of this study. She took all the responsibility of taking care of our children, prayed for my success, and did a lot of *scholarships* for the continuity of my PhD scholarship. She could still find time to cook delicious meals meanwhile. Our children Abia, Abdullah, and Ahmad Bilal deserve special thanks for being a continuous source of energy. They teach me and make my everyday full of life. Without them, this work would have merely been a far dream.

Ziaul Haq Abbas
June 2012
Minneapolis, USA

Abstract

Wireless Internet access is almost pervasive nowadays, and many types of wireless networks can be used to access the Internet. However, along with this growth, there is an even greater concern about the energy consumption and efficiency of mobile devices as well as of the supporting networks, triggering the appearance of the concept of green communication. While some efforts have been made towards this direction, challenges still exist and need to be tackled from diverse perspectives.

Cellular networks, WLANs, and ad hoc networks in the form of wireless mesh networks are the most popular technologies for wireless Internet access. The availability of such a variety of access networks has also paved the way to explore synergistic approaches for Internet access, leading to the concept of hybrid networks and relay communications. In addition, many mobile devices are being equipped with multiple interfaces, enabling them to operate in hybrid networks. In contrast, the improvements in the battery technology itself have not matched the pace of the emerging mobile applications. The situation becomes more sophisticated when a mobile device functions also as a relay node to forward other station's data. In the literature, energy efficiency of mobile devices has been addressed from various perspectives such as protocol-level efforts, battery management efforts, etc. However, there is little work on energy efficiency in hybrid mobile and wireless networks and devices with heterogeneous connections. For example, when there are multiple networks available to a mobile device, how to achieve optimum long-term energy consumption of such a device is an open question.

Furthermore, in today's cellular networks, micro-, pico-, and femto-cells are the most popular network topologies in order to support high data rate services and high user density. With the growth of such small-cell solutions, the energy consumption of these networks is also becoming an important concern for operators. Towards this direction, various solutions have been proposed, ranging from deployment strategies for base stations to cooperative techniques etc. However, as base stations have the largest share in a network's energy consumption, methods that allow lightly-loaded base stations sleep or be switched off are possible means as a feasible step towards green communications.

In this dissertation, we tackle the above mentioned problems from two perspectives, i.e., mobile station's and operator's perspectives. More specifically, by taking into account the amount of transferred data in uplinks and downlinks individually for various components in a hybrid network, strategies are proposed to reduce mobile station's battery energy consumption. For this purpose, other parameters such

as link distance and remaining battery energy can also be considered for handover decision making, in order to maximize energy efficiency of the mobile station. To optimize long-term energy consumption of the mobile stations operated in such scenarios, a Markov decision process-based methodology is proposed as our contribution to this topic. Moreover, from operator's perspective, a network energy conservation scheme which may switch off a base station is proposed for micro- or pico-cells scenarios. Both deterministic and probabilistic schemes are proposed for network energy conservation. The problems considered and the solutions proposed in this dissertation advance the frontiers of the research work within the theme of energy efficiency for mobile devices as well as hybrid mobile and wireless networks.

List of Publications

The purpose of this list is to keep a record of all the scientific publications as the outcome of the research work carried out by the author of this dissertation. The list of publications comprises of one submitted and eight accepted/published papers, organized in two sets, where the first set constitutes Part II of the thesis.

Set I: Papers Included in this Dissertation

The following papers are included in Part II of this dissertation, as Papers A–D respectively.

Paper A Z. H. Abbas and F. Y. Li, A Continuous-space Analytical Approach for Relay Node Placement in Hybrid Cellular and Ad Hoc Networks, in *Proceedings of the 6th IEEE International Symposium on Wireless Communication Systems (ISWCS 2009)*, Siena-Tuscany, Italy, September 2009.

Paper B Z. H. Abbas and F. Y. Li, Analysis of Mobile-oriented Energy Consumption for Heterogeneous Connections in Hybrid Wireless Networks, *International Journal of Communication Networks and Distributed Systems (IJCNDS),* , Inderscience Publishers, vol. 9, no. 3/4, pp. 184–204, 2012.
Available at http://www.inderscience.com/info/jhome.php?jcode=ijcnds

Paper C Z. H. Abbas and F. Y. Li, Two Teletraffic-based Schemes for Energy Saving in Cellular Networks with Micro-cells, *Journal of Communications (JCM)*, Academy Publisher, vol. 7, no. xx, pp. xx–xx, Accepted for publication, 1st June 2012.

Paper D Z. H. Abbas and F. Y. Li, Analysis of Mobile Station Battery Energy Consumption in Heterogeneous Networks using Markov Processes with State Transition Rewards, to be submitted to *IEEE ICC 2013*, Budapest, Hungary, June 2013.

Set II: Papers Not Included in this Dissertation

The papers listed below are complimentary to those listed in Set I. They are, however, not reproduced as a part of the dissertation in order to highlight the most significant contributions of this thesis work.

Paper 5 Z. H. Abbas and F. Y. Li, Power Consumption Analysis for Mobile Stations in Hybrid Relay-assisted Wireless Networks, in *Proceedings of the 5th IEEE International Symposium on Wireless Pervasive Computing 2010 (ISWPC 2010)*, Modena, Italy, May 2010.

Paper 6 Z. H. Abbas and F. Y. Li, Distance-related Energy Consumption Analysis for Mobile/Relay Stations in Heterogeneous Wireless Networks, in *Proceedings of the 6th IEEE International Symposium on Wireless Communication Systems 2010 (ISWCS 2010)*, York, UK, September 2010.

Paper 7 Z. H. Abbas and F. Y. Li, A System-level Power Saving Approach for Cellular Networks with Microcells/Picocells, in *Proceedings of the 2nd IEEE International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace, and Electronic Systems Technology 2011 (Wireless VITAE 2011)*, Chennai, India, February-March 2011.

Paper 8 Z. H. Abbas and F. Y. Li, A Novel Teletraffic-based Power Saving Scheme for Cellular Networks with Microcells, in *Proceedings of the 7th IEEE Performance and Management of Wireless and Mobile Networks (P2MNET) Workshop in conjunction with 36th IEEE LCN*, Bonn, Germany, October 2011.

Paper 9 Z. H. Abbas and F. Y. Li, Energy Optimization in Cellular Networks with Micro-/Pico-cells using Markov Decision Process, in *Proceedings of the 18th European Wireless Conference 2012 (EW 2012)*, Poznań, Poland, April 2012.

Contents

Preface vii

Acknowledgements ix

Abstract xi

List of Publications xiv

List of Figures xxi

List of Tables xxiii

Abbreviations xxv

Part I 1

1 Introduction 3

1.1 Mobile and Wireless Networks 3
1.2 Green Communications and Energy Efficiency 5
1.2.1 Green Communications from the Network Perspective . . . 5
1.2.2 Green Communications from the Mobile Station Perspective 7
1.3 Research Objectives 7
1.4 Organization of the Dissertation 9

2 Relay Node Placement 13

2.1 Introduction to Cellular and Wireless Networks 13
2.1.1 Cellular Networks 14
2.1.2 Wireless Local Area Networks 14
2.1.3 Ad Hoc Networks 15
2.2 Hybrid Networks 15
2.2.1 The Concept of Hybrid Networks 15

2.2.2 Hybrid Network Architectures 16
2.3 Network Coverage Extension 16
2.3.1 Relay Technology for Coverage Extension 16
2.3.2 Capacity and Connectivity in Relay-assisted Networks . . . 17
2.4 Relay Node Placement: The Proposed Approach 18
2.5 Distinguishing Aspects of Paper A 20
2.6 Chapter Summary 20

3 Mobile-oriented Energy Consumption 23
3.1 The Need for Mobile Energy Conservation 23
3.2 Access Mechanisms for IEEE 802.11-based WLANs 24
3.2.1 Distributed Coordination Function 24
3.2.2 Point Coordination Function 25
3.3 Energy-efficient Mobile Networking: A Retrospect 26
3.3.1 Component-oriented Efforts for Mobile Battery Lifetime Extension 26
3.3.2 Mobile-oriented Energy Consumption for Heterogeneous Connections 27
3.4 Analysis of Link Distance-based Mobile-centric Energy Consumption: The Proposed Model 28
3.4.1 Hybrid Network Scenario and Possible Connections 29
3.4.2 Power Consumption in the LDR Cellular Link 30
3.4.3 Power Consumption in the HDR Ad Hoc Link 30
3.4.4 Power Consumption in the HDR WLAN Link 31
3.4.5 Major Contributions of Paper B 32
3.5 Chapter Summary 33

4 Towards Green Communications in Cellular Networks 35
4.1 Energy Consumption in Cellular Networks: An Overview 35
4.1.1 Energy Consumption from Operators' Perspective 35
4.1.2 Energy Consumption from MSs' Perspective 36
4.2 Energy Reduction Approaches in Cellular Networks 37
4.2.1 Architecture-level Energy Saving in Cellular Networks . . . 38
4.2.2 Energy Conservation through Emerging Technologies . . . 39
4.2.3 Network Planning for Energy Conservation 40
4.3 A Deterministic Approach for Power Saving 41
4.4 A Probabilistic Approach for Power Saving 44
4.5 Distinguishing Aspects of Paper C 45

4.6 Chapter Summary 46

5 Markov Decision Process for Battery Energy Optimization 47

5.1 Introduction to Mobile Battery Consumption Estimation 47

5.2 A Retrospect on Energy Management Techniques for Mobile Devices 48

5.2.1 Protocol-level View 48

5.2.2 Measurements and Models for Energy Consumption 49

5.2.3 MS Energy Consumption Prediction and Optimization . . . 50

5.3 Energy Consumption Prediction and Optimization in Heterogeneous Networks: Our Approach 50

5.3.1 Discrete-Time Markov Chain-based Energy Prediction . . . 51

5.3.2 Finite Markov Decision Process for Energy Optimization . . 52

5.4 Distinguishing Aspects of Paper D 53

5.5 Chapter Summary 53

6 Concluding Remarks 55

6.1 Summary of the Research Work and Major Scientific Contributions 55

6.2 Limitations of the Research 57

6.3 Suggestions for Future Research 57

References 59

Part II 73

A A Continuous-space Analytical Approach for Relay Node Placement in Hybrid Cellular and Ad Hoc Networks 75

B Analysis of Mobile-oriented Energy Consumption for Heterogeneous Connections in Hybrid Wireless Networks 89

C Two Teletraffic-based Schemes for Energy Saving in Cellular Networks with Micro-cells 117

D Analysis of Mobile Station Battery Energy Consumption in Heterogeneous Networks using Markov Processes with State Transition Rewards 149

List of Figures

1.1 Layout of the goals and achievements in this dissertation. 11

2.1 Connecting conventional phones with cellular mobile phones. 14
2.2 Extension of cellular coverage using an ad hoc network. 16
2.3 Illustration of the regions in the HN. 19

3.1 CSMA/CA mechanism illustration. 24
3.2 A hybrid network scenario with three heterogeneous links. 29
3.3 Illustration of contention free repetition interval. 31
3.4 Interaction between transmitter and receiver during a CP. 31
3.5 Interaction between transmitter and receiver during CFP. 32

4.1 Network scenario used in Paper C, where the BS in Region-Z can be switched off under light traffic conditions. 41
4.2 Power saving policy used in Paper C, where a hysteresis region is introduced to avoid the ping-pong effect. 42
4.3 Total transmission consumptions: (a) $\lambda/\mu = 0.5$, $t_l + t'_l = 7$, $t_u + t'_u = 12$, (b) $\lambda/\mu = 0.5$, $t_l + t'_l = 6$, $t_u + t'_u = 10$. 44

5.1 Partial FMDP for the MS with initial state at W. 52

A.1 Hybrid network architecture. 79
A.2 Basic model for total interference calculation (ad hoc to relay). . . . 81
A.3 Variation of r and placement of relay node. 83
A.4 Capacity curves for the hybrid network with lower path-loss constant: (a) Sparsely populated, (b) Moderately populated, (c) Densely populated. 86
A.5 Capacity curves for the hybrid network with higher path-loss constant: (a) Sparsely populated, (b) Moderately populated, (c) Densely populated. 87

B.1 Basic heterogeneous network architecture. 95

B.2 Cellular link energy consumption: effects of α and distance. 99
B.3 Division of CFRI into CFP and CP. 100
B.4 Interaction between transmitter and receiver during CP. 101
B.5 Subdivision of CFP. 102
B.6 Direct cellular link: relation between remaining energy of MS and (a) bytes exchanged (b) battery lifetime. 106
B.7 WLAN link: relation between remaining energy of MS and (a) bytes exchanged (b) battery lifetime. 107
B.8 URDR connection: relation between remaining energy of MS and RS (a) bytes exchanged (b) battery lifetime. 108
B.9 UCDW and UWDC connections: relation between remaining energy of MS (a) bytes exchanged (b) battery lifetime 109
B.10 URDW and UWDR connections: relation between remaining energy of MS (a) bytes exchanged (b) battery lifetime. 110
B.11 URDW and UWDR connections: relation between remaining energy of RS (a) bytes exchanged (b) battery lifetime 111
B.12 URDW and UWDR connections: relation between remaining energy of RS and bytes exchanged for asymmetric traffic 112

C.1 Network scenario (a) Overall network; (b) Definitions of regions. . . 122
C.2 BD process for Region-W. 124
C.3 BD process for Region-C. 125
C.4 Power saving policy for the network. 126
C.5 Effects of traffic load, A, on the probability of channel occupancy. . 129
C.6 Deterministic scheme: Total transmission and network power consumptions for (a) $\lambda/\mu = 0.4$, $t_l + t'_l = 5$, $t_u + t'_u = 10$; (b) $\lambda/\mu = 0.6$, $t_l + t'_l = 5$, $t_u + t'_u = 10$. 137
C.7 Deterministic scheme: Total transmission and network power consumptions for (a) $\lambda/\mu = 0.4$, $t_l + t'_l = 6$, $t_u + t'_u = 11$; (b) $\lambda/\mu = 0.6$, $t_l + t'_l = 6$, $t_u + t'_u = 11$. 139
C.8 Probabilistic scheme: Non-optimal network transmission and total power consumption for (a) $\lambda/\mu = 0.4$, and (b) $\lambda/\mu = 0.6$. 141
C.9 Optimum network transmission power consumption. 142
C.10 Difference between non-optimal and optimum transmission power consumption in the PS-ON mode (upper part) and the PS-OFF mode (lower part). 142
C.11 **P**-flipped: Optimum network transmission power consumption. . . . 143

C.12 **P**-flipped: Difference between non-optimal and optimum transmission power consumption in the PS-ON mode (upper part) and the PS-OFF mode (lower part). 144

D.1 Network scenario for MS with heterogeneous connections. 153
D.2 Partial FMDPs for the MS: initial state at (a) W, (b) A, and (c) C. . . . 159
D.3 Energy consumption using DTMC when the MS starts in state W, A, or C, for various distances between the MS and the BS, AP, or RS. 161
D.4 Minimized energy consumption using FMDP for various distances between the MS and the BS, AP, or RS. 163
D.5 Minimized energy consumption using FMDP for various distances between the MS and the BS, AP, or RS. 163

List of Tables

1.1 Mapping of research goals and research questions. 9

B.1 Ad hoc and WLAN link analysis parameters. 105

C.1 Analysis parameters. 136

D.1 Parameter configuration. 160

Abbreviations

1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	Third Generation Partnership Project
3GPP2	Third Generation Partnership Project Two
4G	Fourth Generation
A/V	Audio/Video
AC	Alternating Current
ACK	Acknowledgment
AP	Access Point
AWGN	Additive White Gaussian Noise
B3G	Beyond 3G
BS	Base Station
BSS	Basic Service Set
CAPEX	Capital Expenditure
CDMA	Code Division Multiple Access
CFP	Contention-Free Period
CFRI	Contention-Free Repetition Interval
CO_2	Carbon Dioxide
CP	Contention Period
CPU	Central Processing Unit
CSMA/CA	Carrier Sense Multiple Access/Collision Avoidance
CTS	Clear To Send
DCF	Distributed Coordination Function
DIFS	DCF Inter Frame Spacing
DRX	Discontinuous Reception
DS	Distribution System
DTMC	Discrete-Time Markov Chain
DTX	Discontinuous Transmission

EC	European Commission
EI	Energy Ignorant
ESS	Extended Service Set
FACH	Forward Access Channel
FDMA	Frequency Division Multiple Access
FMDP	Finite Markov Decision Process
GBS	Green Base Station
GoS	Grade of Service
GSM	Global System for Mobile communications
HDR	High Data Rate
HES	High Energy Saving
HN	Hybrid Network
IBSS	Independent Basic Service Set
ICAR	Integrated Cellular and Ad hoc Relaying
ICE	Intelligent Cell brEathing
ICT	Information and Communication Technology
IEEE	Institute of Electrical and Electronics Engineers
IP	Internet Protocol
ISDN	Integrated Services Digital Network
LDR	Low Data Rate
LES	Low Energy Saving
LTE	Long Term Evolution
LTE-A	Long Term Evolution-Advanced
MAC	Medium Access Control
MC	Markov Chain
MIH	Media Independent Handover
MS	Mobile Station
MSC	Mobile Switching Centre
NAV	Network Allocation Vector
NIC	Network Interface Card
OPEX	Operational Expenditure
PA	Power Amplifier
PC	Personal Computer
PCF	Point Coordination Function
PSM	Power Saving Mode
PS-OFF	Power Save OFF
PS-ON	Power Save ON

PSTN	Public Switched Telephone Network
QoS	Quality of Service
RF	Radio Frequency
RNC	Radio Network Controller
RS	Relay Station
RTS	Request To Send
SIFS	Short Inter-Frame Spacing
SIR	Signal-to-Interference Ratio
SON	Self Organizing Network
STA	Station
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
UCAN	Unified Cellular and Ad hoc Network
UCDW	Uplink Cellular Downlink WLAN
UMTS	Universal Mobile Telecommunication System
URDR	Uplink Relayed Downlink Relayed
URDW	Uplink Relayed Downlink WLAN
UWDC	Uplink WLAN Downlink Cellular
UWDR	Uplink WLAN Downlink Relayed
VoIP	Voice over Internet Protocol
WiFi	Wireless Fidelity
WiMAX	Worldwide interoperability for Microwave Access
WLAN	Wireless Local Area Network

Part I

Chapter 1

Introduction

1.1 Mobile and Wireless Networks

The mobile market has grown dramatically during the past two decades, surpassing the number of wireline subscribers since 2001. At the same time, mobile phones have become much more than just communication devices. With the popularity of 3rd Generation (3G) mobile communication technology, mobiles devices are now multimedia-enabled, delivering a variety of applications including music, games, video-streaming, web-browsing, and e-mail among others. At the same time, the wireless industry is also experiencing explosive and sustainable growth. The drivers for this growth, although different for each network type, are owing to ubiquitous network coverage, easy and cheap access, as well as emerging applications.

Indeed, mobile communication is the fastest growing and most demanding sector in the telecommunications world. New generations of cellular systems, which offer accessibility and versatility, and mobile handsets that support a range of innovative services, are the design objectives and the main interests of not only telecommunication industry, network operators and service providers, but also the research community. Furthermore, while the surge in user applications is only bounded by imagination, the capacity of cellular systems, in order to support these applications, is tightly restricted by fundamental physical limits. With the increasing demand in the ubiquity of cellular networks, the number of infrastructure equipment is rapidly increasing in the meantime. The problem is further aggravated by considering financial constraints from the operator's point of view, as higher capacity and better Quality of Service (QoS) come at the cost of higher Capital Expenditure (CAPEX) and Operating Expenditure (OPEX). Since users may be reluctant to pay proportionally higher bills for improved services, minimizing CAPEX and OPEX in order to make the business model commercially viable, whilst seeking to provide better

QoS and higher capacity, is a crucial requirement for today's operators [1].

On the one hand, users' interests and demands for better services are the driving force for the evolution of new technology and the enhancements based on existing technologies. The ever-increasing demand of mobile users for wireless access of multimedia services leads to challenging tasks like providing both higher data rate and better coverage, as well as reducing data transfer cost. Wireless Local Area Networks (WLANs), represented by IEEE 802.11, provide such a niche between user demands and network costs. On the other hand, Beyond 3G (B3G) mobile communication technologies allow mobile devices to access the Internet via multiple interfaces. In [2], a summary of diverse existing wireless and wired access technologies is given, including WLAN, Wireless Fidelity (WiFi), Worldwide Interoperability for Microwave Access (WiMAX); cellular technologies such as Global System for Mobile communications (GSM), General Packet Radio Service (GPRS), Universal Mobile Telecommunication System (UMTS); and other earlier generation wired networks such as Public Switched Telephone Network (PSTN), Integrated Services Digital Network (ISDN) and so on. These networks provide different types of services with different coverage, data rates, cost etc to end users.

At the same time, the research work in the field of wireless ad hoc networks, which targets at allowing wireless hosts communicate with each other without relying on fixed infrastructure, is gaining more and more momentum [3]. Examples of possible applications of ad hoc networking include municipality wireless networks, emergency disaster relief, networks of mobile devices, etc. When an ad hoc network is created, multiple hops may be required for a source to reach its destination. Hence, relay nodes are a fundamental requirement for multi-hop ad hoc networks.

Furthermore, to explore the effects of synergistic approaches, the interest of the research community in the field of wireless Hybrid Networks (HNs), sometimes also referred to as Heterogeneous Networks, is increasing. HNs are the mobile and wireless networks that are formed by a collection of wireless Access Points (APs) and stations, nodes in ad hoc networks, cellular networks, and wireless enabled computing devices etc. Moreover, with the rising popularity of HNs and multimedia-enabled mobile devices, the energy conservation aspect of the constituent networks and mobile devices has become a topic of great interest. In this dissertation, we address a few issues within the topic of energy efficiency in hybrid mobile and wireless networks.

1.2 Green Communications and Energy Efficiency

As the number of cellular and wireless networks as well as the number of mobile users explode, energy efficiency has become a major concern. Indeed, the energy consumption problem in the Information and Communication Technology (ICT) sector has become crucial during the past years. On the one hand, ICT is expected to play a key role in reducing the energy consumption in many sectors such as transportation, power, agriculture, etc., which are the major contributors to the rise of global Carbon DiOxide (CO_2) emission. As an instance, a recent study [4] estimates that ICT can reduce up to 25 percent energy consumption in the transport sector and about 30 percent in the manufacturing sector. Moreover, ICT is expected to significantly improve the efficiency of energy generation and distribution through the concept of smart grid. On the other hand, ICT especially mobile industry itself is also a contributor of CO_2 emission through network operations, mobile equipments etc. According to [5], ICT equipment is responsible for about $2 \sim 10$ percent of the world energy consumption. Indeed, these facts have attracted a keen interest among the research community in the field of energy efficient ICT, triggering the appearance of a popular terminology – *green communication*. Furthermore, due to the above mentioned reasons, there is now a worldwide effort towards energy efficient solutions, evidenced through several large-scale initiatives [6, 7, 8, 9].

More specifically, green communication is an innovative research area to find radio communication and networking solutions that can greatly improve energy efficiency as well as resource efficiency of wireless communications without compromising the QoS of users. It not only contributes to global environment improvement but also achieves commercial benefits for telecommunication operators. To meet the challenges of increasing energy efficiency in communication systems, it is imperative to resort to paradigm-shifting technologies, such as energy efficient network architectures, energy efficient wireless transmission techniques, energy efficient networks and protocols, smart grids, etc. Some recent efforts towards achieving green communication solutions include [10, 11, 12, 13]. Furthermore, in order to achieve real green wireless and cellular communications, the energy efficiency of both networks and mobile devices needs to be addressed evenly.

1.2.1 Green Communications from the Network Perspective

The mobile industry is facing a critical energy conservation challenge. By year 2013 the number of smartphones will exceed 1.82 billions worldwide and will surpass PCs as the most popular web access devices. This fact has huge impact on the amount of energy consumed by the supporting infrastructure equipment. Meanwhile, the network data volume is expected to increase by a factor of 10 every five years, associated with about 20 percent increase of energy consumption correspondingly. Therefore, the mobile industry faces a great sustainable development problem in energy consumption.

Until recently, the research and development efforts in the area of communication networks have been mainly targeted at functionality and performance. For battery-operated devices alone, such as Mobile Stations (MSs) and laptops, was energy efficiency a significant consideration. This trend has however significantly changed with the massive use of wireless communications in public, professional, and private lives coupled with the increasing cost of energy. From a cellular network operator's perspective, reducing energy consumption can also be translated into lower OPEX. As Base Stations (BSs) are the most energy consuming constituent of a cellular network [14], they have become a highly actual candidate for the adoption of energy efficient approaches in many networking scenarios. This is evidenced by many efforts towards energy-efficient cellular network solutions [15, 16, 17, 18, 19]. More such efforts are discussed in later chapters of this thesis. Similarly, APs in WLANs play the role as a major energy consumer, especially in dense environments. Therefore, efficient energy management of mobile and wireless networks becomes of paramount importance for green communications [20, 21].

Furthermore, cell sites are tending to get smaller nowadays in order to support high device density and multimedia services. While giant macro-cell sites have been the industry norm for years, operators are now increasingly deploying a range of smaller cell sites as a way to reduce cost and speed up network expansion to meet the growing demands on high data rate and user population. These cells include micro-, pico-, and femto-cells. Micro-cells, frequently deployed in urban and suburban areas, offer a coverage radius of less than one kilometer in diameter. Pico-cells are even smaller. They cover areas of a few hundred meters in radius and are typically used for indoor applications such as office buildings, shopping malls etc. Femto-cells are the smallest of the bunch, and their coverage area is just a few meters targeting at office and home scenarios. With the increasing popularity of small cell site solutions, the number of corresponding infrastructure is also expanding. Therefore, their energy consumption is rapidly becoming a concern in operators' community.

1.2.2 Green Communications from the Mobile Station Perspective

Studies indicate that the power drain for an MS per subscriber is much lower than that for the BS [14]. However, the effect of MS energy consumption cannot be ignored, neither from the individual user's point of view nor from the environmental perspective. Moreover, despite the efforts in recent years, the reduction of carbon footprints from the MS usage perspective has not been adequately explored. Most of the existing work is targeted at reducing power consumption by incorporating smart mobile applications or invoking standby mode in MSs. There are a few approaches attempting to optimize the energy efficiency of MSs from Medium Access Control (MAC) protocol aspects of the IEEE 802.11 WLAN [22] [23]. Another recent investigation on the combined performance of both power saving and QoS features in IEEE 802.11e based WLANs can be found in [24]. In another attempt [25], the authors analyze the impact of network discovery on the MS's power consumption through network scanning and broadcasting approaches and report about $8 \sim 35$ percent energy saving for the involved MSs.

In addition, with the popularity of heterogeneous networks, MSs are being equipped with multiple interfaces. Therefore, MSs can also be used to bridge more than one networks, e.g., through relaying techniques. This trend has led to further complexity in the study of MS energy consumption reduction/optimization operated in HNs.

Considering above background information, the focus of this dissertation has been twofold: 1) energy consumption reduction of MSs operated in HNs; and 2) network energy conservation.

1.3 Research Objectives

The objective of this dissertation is twofold. First, to investigate the mechanisms and approaches to calculate, reduce, and optimize energy consumption of MS battery operated in an HN. The distinct features of HNs have brought us to encounter many interesting problems around MS battery energy consumption. Furthermore, considering the ubiquity of cellular networks and the popularity of small-cell networks such as micro-cells, the second objective is to propose techniques to reduce the energy consumption of cellular networks from operator's perspective. More specifically, we attempt to answer the following research questions in this dissertation:

- Question 1: When the coverage of a cellular network is extended by using relay nodes, where is the optimum position to place such a relay node in the resulted HN, in order to achieve optimum overall capacity?

- Question 2: By calculating the battery lifetime and the amount of transferred bytes in each link of an HN, can we improve the energy efficiency of an MS operated in the HN by utilizing different paths for uplink and downlink traffic?

- Question 3: In a network with micro- or pico-cells, how can we reduce overall network energy consumption according to traffic variations? Furthermore, can we obtain an optimum long-term figure of the overall network energy consumption for such networks?

- Question 4: In an HN, when an MS can probabilistically switch access across the available constituent networks, how can we achieve long-term optimum energy consumption by analyzing the battery energy consumption of such an MS?

Based on the above research questions, a detailed literature review was performed to survey existing solutions and investigate potential technologies and approaches. Correspondingly, the following research goals were identified and pursued:

- Goal 1: To explore the optimum capacity in an HN consisting of a cellular network which is extended through the usage of relay nodes from an ad hoc network. Furthermore, for given node density and path-loss coefficient, to determine the precise location for relay node placement in order to achieve maximum overall capacity.

- Goal 2: To propose an MS energy conservation approach by using distance-based energy consumption figures for various links within an HN. Furthermore, to investigate the possibilities of reducing MS energy consumption for mobile-centric energy-aware handover in HNs.

- Goal 3: To develop teletraffic based schemes for energy conservation in a cellular network with micro- or pico-cells by exploring the idea of switching-off BS through power-saving policy development. Furthermore, to propose energy conservation schemes for such networks and the proposed power-saving policy, which is in contrast with many existing techniques, should be based on traffic intensity, not time of a day.

- Goal 4: To analyze MS battery energy consumption in HNs using Markov processes with state transition rewards in order to obtain long-term as well as optimum energy consumption figures.

Table 1.1 illustrates the mapping between the targeted research goals and the addressed research questions. The goals are achieved through a set of scientific contributions which include Part I of this dissertation and Papers A-D in Part II. Each of the chapters from Chapter 2 to Chapter 5 in this thesis gives a survey on the relevant literature and outlines the scientific contributions of the thesis work within each identified topic.

Table 1.1: Mapping of research goals and research questions.

Research Goal	**Research Question**
Goal 1	Question 1
Goal 2	Question 2
Goal 3	Question 3
Goal 4	Question 4

1.4 Organization of the Dissertation

This dissertation is partitioned into two parts. Part I provides an overview of the PhD work and consists of Chapters 1-6. Part II reprints four scientific publications, Papers A-D. Fig. 1.1 illustrates the logical connections between different topics and papers around which the whole dissertation is built up.

The rest of Part I of this dissertation is organized as follows:

- *Chapter 2* presents an overview of cellular, WLAN, and ad hoc networks, and introduces the concept of hybrid networks. The problem of relay node placement to extend cellular network coverage is then discussed in this chapter. We propose a continuous-space analytical approach based on circular geometry of the network region for relay node placement in such HNs.

- *Chapter 3* sheds light on energy consumption of MSs in HNs. The state-of-the-art efforts on energy conservation for MSs are summarized. We propose a novel link-distance based mobile-centric battery energy conservation methodology for an MS operated in HNs.

- *Chapter 4* provides a detailed literature survey on the issues of network energy conservation for cellular networks. Energy reduction techniques from the perspectives of network architecture, enabling technologies and network planning are summarized. Later on, we propose two approaches for network energy conservation in a cellular network with micro- or pico-cells, and analyze their performance.

- *Chapter 5* is devoted to battery energy prediction and optimization for an MS operated in an HN. A Markov analysis based methodology for MS energy conservation in such networks is performed. We further propose a Finite Markov Decision Process (FMDP) based methodology for long-term optimum MS energy consumption prediction.

- *Chapter 6* recapitulates the major scientific contributions of this thesis work and presents the limitations of the current research, then points out a few future research directions.

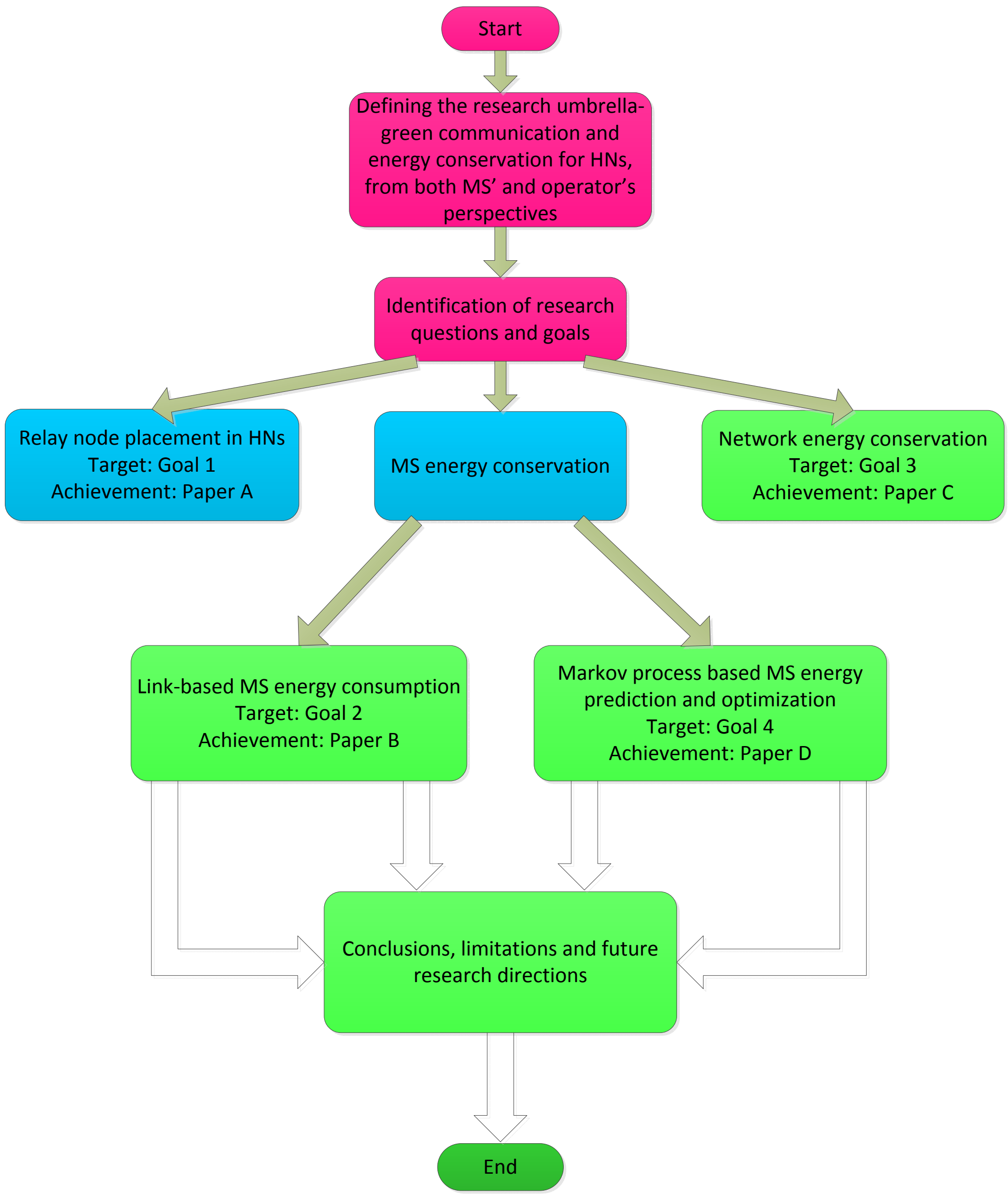


Figure 1.1: Layout of the goals and achievements in this dissertation.

Chapter 2

Relay Node Placement in Hybrid Cellular and Ad Hoc Networks

2.1 Introduction to Cellular and Wireless Networks

Mobile communication systems have revolutionized the way people live and communicate. The evolution of mobile communication technologies is now reaching its 4th generation (4G). While the first generation (1G) fulfilled the voice requirements and was based on analogue technology, the second generation (2G) was based on digital technology and targeted at achieving higher capacity in mobile communications systems. Text messaging was also introduced in 2G. The third generation offers multimedia services and Internet connectivity to mobile phones. The fourth generation targets towards an all-IP solution and promises to achieve ultra high data rate. At the same time, various types of access networks, ranging from cellular to WLANs and from short-range communication to relay-based networks, are supported by many operators.

On the other hand, the evolution of networking technologies has also revolutionized the capabilities of mobile phones. Being originally designed for simple voice and messaging services, mobile phones have now reached a stage in which they can be compared to medium-scale PCs. Furthermore, thanks to the advances in both networking and mobile communication technologies, mobile devices are able to connect to the Internet via heterogeneous links, e.g., through cellular systems, WLANs, and/or ad hoc networks.

2.1.1 Cellular Networks

As an extension to fixed telephony, a cellular network allows user mobility and integrates cellular phones or MSs into the conventional PSTN. The service coverage region of a cellular network is divided into many small or large areas, known as cells. Each cell is served by a BS in 2G or Node B in 3G. The BS/Node B is fixed and connected to a Mobile Switching Center (MSC) in 2G or a Radio Network Controller (RNC) in 3G. An MSC or RNC is in charge of a cluster of BSs, and through them, MSs are connected to the PSTN and the Internet. Both MSs and BSs are equipped with a transceiver. Hexagonal geometry is a conventional pictorial representation of cellular networks. Fig. 2.1 illustrates a cellular network connected to a conventional phone.

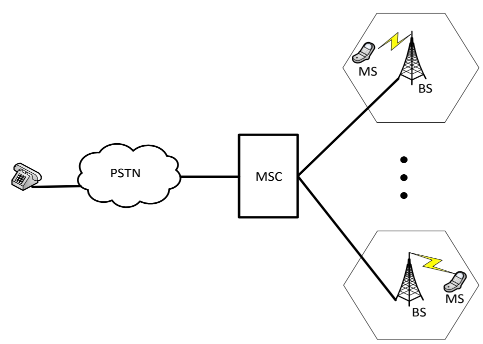


Figure 2.1: Connecting conventional phones with cellular mobile phones.

The success of cellular networks is primarily due to the concepts of cell and frequency reuse, where each BS (or cell) is assigned a set of frequency bands. Within a cell, different users can access the channel through either Frequency Division Multiple Access (FDMA), Time Division Multiple Access (TDMA), or Code Division Multiple Access (CDMA).

2.1.2 Wireless Local Area Networks

The major components in a WLAN architecture include Station (STA), AP, Basic Service Set (BSS), Extended Service Set (ESS), and Independent Basic Service Set (IBSS). WLANs can operate in two modes, i.e., infrastructure mode and ad hoc mode. In the infrastructure mode, stations are associated with an AP then connected to the Internet. All traffic among STAs or between a STA and the Internet has to go via the AP in this case. In an ESS, there are more than one APs and these APs are connected with each other through a Distribution System (DS). In the ad hoc mode, the network operates as an IBSS which allows peer-to-peer direct communications among STAs. The most popular WLAN implementation and deployment are based on the IEEE 802.11 standard.

2.1.3 Ad Hoc Networks

An ad hoc network is a transient network formed dynamically by a collection of arbitrarily located wireless devices, without the necessity of relying on any existing network infrastructure [3]. This may also be referred to as spontaneous networking [26]. Originally ad hoc networks are targeted at tactical network oriented applications to improve communications and survivability in a battlefield. In recent years, ad hoc networking technologies have been successfully deployed in many civil applications like wireless mesh networks for Internet access.

In an ad hoc network, the devices themselves form the network and exchange information among them, either directly or via multiple hops. That is, nodes cooperate with each other to provide networking functionality, as both hosts and routers. In such a network, nearby terminals communicate with each other directly. The devices that cannot communicate with each other directly may get their traffic forwarded through one or more intermediate devices, referred to as relay nodes. The most popular enabling one-hop technology for building multi-hop ad hoc networks is again based on the IEEE 802.11 standard.

2.2 Hybrid Networks

2.2.1 The Concept of Hybrid Networks

A hybrid network is created when the features of one or more networks are used to improve certain characteristics of another network. It is composed of a collection of all the constituent networks' elements. One of the major reasons that ad hoc networks gained momentum is because they provide networking capability in areas where no (or limited) communication infrastructure pre-exists, or when the coverage of such infrastructure requires wireless extension [27] [28]. Furthermore, in situations where there is a need for cellular coverage (or capacity) extension, ad hoc networks provide an attractive low-cost alternative. In this context, an HN created by using relay nodes to extend cellular network range appears as a feasible solution. As an example, Fig. 2.2 illustrates an HN where the cellular coverage extension can be achieved with the help of relay nodes through ad hoc networking. In such an HN, the BS may reduce its coverage area to have direct communications with only a subset of users within each cell. The users that are located outside the direct

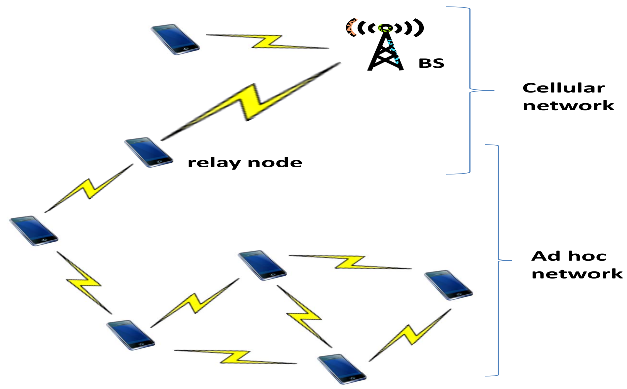


Figure 2.2: Extension of cellular coverage using an ad hoc network.

coverage area receive packets to/from the BS via relays, in a two-hop or multi-hop fashion. The relay node covers short distance but with higher data rate.

In an HN, the network capacity is not a function of just one network, but of all the constituent networks, as the maximum possible bits per second under the conditions of max-min fairness [29]. For a given set of parameters, the computed capacity may or may not be achievable in realistic systems. However, the achievable throughput is always lower than or equal to the theoretical capacity.

2.2.2 Hybrid Network Architectures

A few HN models have been proposed in the literature. An integrated Cellular and Ad hoc Relaying (iCAR) system is proposed in [30]. In this model, ad hoc relay stations are deployed strategically to improve call service quality, but capacity gain is not the main focus of their paper. The authors in [31] propose a Unified Cellular and Ad hoc Network (UCAN) architecture to improve individual user throughput and the aggregate downlink throughput. UCAN assumes the use of an additional spectral band for ad hoc communications.

Moreover in [32], the authors consider an ad hoc network model for wireless packet data network. It is shown that the network can support better spatial reuse and result in better throughput per unit power spent.

2.3 Network Coverage Extension

2.3.1 Relay Technology for Coverage Extension

As already mentioned, relay technology can be used for network coverage extension and capacity enhancement due to its simplicity, flexibility, ease of deployment, and cost effectiveness. A relay node is an intermediate network node placed between the source and the destination nodes in order to help forwarding the information sent by the source node to combat capacity degradation or packet loss. As it is not possible to achieve high data rate over long distance, intermediate relay nodes can be used to improve such situations. In fact, relaying is one of the proposed technologies for Long Term Evolution-Advanced (LTE-A) networks [33] for coverage extension with high data rate. Relay nodes in such scenarios do not necessarily need to be connected through wires, like copper or fiber, thereby offering high flexibility for relay placement, fast network rollout and adaptive traffic capacity engineering [34]. Other notable attributes of relay-based networks include alternative multiple routes, optimized network load distribution etc.

Furthermore, in areas like rural regions where traffic density is low and the population is sparsely distributed, it becomes less viable to build traditional cellular access networks with full BS deployments. In such cases, a network architecture with a single BS assisted by relay nodes (either mobile or stationary) to improve capacity and coverage may be a more efficient solution.

2.3.2 Capacity and Connectivity in Relay-assisted Networks

Various attempts exist in order to increase the cellular range and capacity. However, there are also fundamental limits on these approaches. In addition to the relationship between data rate and distance, the transmission power of BSs and MSs cannot be increased beyond pre-defined limits due to safety regulations, interference, and battery life considerations etc. Relay transmission appears as a cost-effective approach in such context.

Moreover, although there has been a surge of interest in modeling various kinds of HNs [30, 31, 32], [35, 36, 37, 38, 39, 40, 41, 42, 43], the question on whether using multi-hop ad hoc wireless relays is beneficial in terms of capacity enhancement of the original cellular network has not been thoroughly investigated yet. While the use of wireless relay nodes may improve spatial and time diversity [44], the employment of multi-hop relays increases the number of hops traversed, driving down the achievable throughput. Given these two conflicting factors, whether the capacity of the network will increase or not in the end with respect to the original cellular network is unclear.

In [45], the authors study the single-tier constrained relay node placement problems, under both the connectivity and the survivability requirements. The relay nodes can only be placed at a subset of candidate locations. The authors discuss the complexity of the problem and propose approximation algorithms to solve the problems. However, their main focus is merely on the connectivity of sensor networks. How the capacity of the network is affected using their schemes is not discussed. The authors in [46] evaluate an HN by reducing the BS coverage and relaying the traffic outside the BS coverage area with the intention to explore whether their scheme could improve the total capacity of the system. Through regular placement of BS and users with hexagonal geometry in the network, they show that moderate gains in the achievable downlink throughput are possible by using an HN, as compared with a pure cellular network. However, how to find the optimum capacity of the HN has not been discussed in their paper. Furthermore, in [47], the impact of mobile relays on throughput and delay is studied by assuming hexagonal grid around the BS, and that the MSs are randomly distributed in a certain region. However, in their work, the relay nodes are pre-selected as the ones which are closest to the BS.

In a seminal paper, [48], the authors analyze the fundamental scaling laws of the maximum achievable network capacity in the presence of heterogeneous nodes and under different mobility models. For an HN with heterogeneous links, a natural question arises as – what will be the picture of the overall capacity of the network and how does the relay node placement affect this overall capacity?

Another question that comes into mind, when a cellular network can be assisted by mobile relays, is how can we trade off the low coverage but high capacity of the ad hoc network with the high coverage but low capacity of the cellular network. Furthermore, another relevant question is how to place the relay nodes such that the overall HN capacity is optimized. In Paper A, we attempt to address these issues.

2.4 Relay Node Placement: The Proposed Approach

In Paper A, which serves as the first endeavour of this thesis work, we analyze the overall capacity of an HN consisting of a cellular and an ad hoc network. The overall network is divided into three concentric circular regions, referred to as A, AB, and B. Region-A is the innermost circular area. Region-B consists of a circular ring of radius greater than radius of Region-A. Region-AB is a circular ring whose outer boundary is smaller than that of outer boundary of Region-B and whose inner boundary is greater than the outer boundary of Region-A. The BS is considered to be placed at the center of these circles. The MSs in Region-A can reach the BS directly, and are referred to as *cellular nodes*. An MS in Region-B is also reachable by the BS, but at a lower data rate. However, if an MS within Region-AB is employed as a relay node, the MSs in Region-B can be reached at a high data rate. Therefore, the

MSs in Region-AB are referred to as *relay nodes*. An illustration of these regions is depicted in Fig. 2.3.

All the MSs in the network are considered to be uniformly distributed. The algorithm for network topology update information is assumed to be working on the backend. All MSs have identical transmission power so that the Signal-to-Interference Ratio (SIR) is mainly dependent on the distance and the path-loss coefficient.

The main task of Paper A is to investigate the optimum overall capacity of the HN, and then to find the optimum range of distance from the BS to place relay nodes. For this purpose, we propose a continuous-space method and divide the overall region into elemental circular subregions (rings) with very small width. These rings serve as the basis of our continuous-space analysis of the overall region.

Initially, the total interference experienced by a receiving MS in the ad hoc mode from all other MSs is calculated by integrating all the individual interference contributions over all the interference region of the MS. The received signal strength at the MS from a relay node is easily obtained by using the relationship between the transmitted and the received power. Dividing the received signal strength at the MS by the total interference experienced by the MS, we obtain the value for SIR. This enables us to determine the ad hoc capacity for given channel bandwidth by using the well-known Shannon formula.

The capacity between the relay node and BS is obtained through similar methodology however with a constraint that there should be at least two stations in the region under consideration.

The total capacity will decrease as the distance between MS–RS transceivers is increased for the ad hoc network part, but increase as the MS moves closer to the BS for the cellular part, given that the total distance from the MS via the relay node to the BS is fixed. Therefore, there will be a point in the HN where the overall total

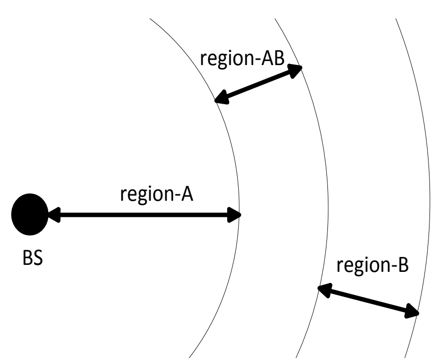


Figure 2.3: Illustration of the regions in the HN.

network capacity is optimized. This means that there exists a point where the plots of these two capacity curves intersect.

In Paper A, the task of placing the relay node in an optimum location is formulated as

$$\begin{aligned} \textit{maximize} \quad & C_{total}^{opt} \\ \text{s.t.} \quad & d_{MS-relay} + d_{relay-BS} = D, \end{aligned} \tag{1}$$

where C_{total}^{opt} is the total optimum network capacity, $d_{MS-relay}$ denotes the distance between the MS and its relay node, $d_{relay-BS}$ is the distance between the relay node and the BS, and D represents the distance from the BS to the boundary of the HN where the MS is located. In other words, we can write the above problem as $C_{total}^{opt} = max\{min(C_{adhoc}, C_{cellular})\}$, where C_{adhoc} and $C_{cellular}$ denote the ad hoc and cellular capacity respectively .

The numerical results in Paper A indicate that we need to place the relay node in a range which is neither too close nor too far away from the BS in order to achieve optimum overall capacity for the HN. Furthermore, for given values of node density and path-loss coefficient, we can find a precise location for relay node placement to achieve maximum overall capacity.

2.5 Distinguishing Aspects of Paper A

Paper A distinguishes from the existing literature in a way that we have proposed a *continuous-space* analytical approach with circular geometry to analyze the network capacity in order to find a position for an optimal relay node placement, whereas most of the other related work is based on a discrete-space approach. We also studied, in this work, how to find this overall capacity with respect to the change in node density and path-loss coefficient.

2.6 Chapter Summary

In an HN, relay nodes, which are the nodes that from an ad hoc network, can be used to extend cellular coverage and compensate capacity decay at the cellular cell edges. However, optimum relay node placement in HNs is a challenging task. This chapter began with an overview of cellular, ad hoc, and hybrid networks. Network coverage extension problems and relevant existing work were examined afterwards. Then we briefly outlined the work carried out in Paper A of the thesis which proposed a continuous-space analytical approach for cellular coverage extension through a relay node. The results demonstrate that for given network topology and parameter values, optimal positions for relay node placement in HNs exist in order to obtain maximum overall network capacity and these positions can be obtained using our proposed continuous-space analysis.

Chapter 3

Mobile-oriented Energy Consumption in Hybrid Networks

3.1 The Need for Mobile Energy Conservation

To access the Internet through wireless connections, cellular systems and WLANs are two most popular alternatives. From energy consumption point of view, the potentially large distance between a mobile phone and its BS leads to higher cost for signal transmissions. On the other hand, shorter-range wireless networks are also becoming increasingly popular these days, especially in urban, residential, and business scenarios. With relatively short range, WiFi achieves much higher data rates however with much limited coverage compared with cellular networks. Therefore, to achieve energy efficient ubiquitous wireless connectivity, it is important to explore the strength of both networks.

However, limited battery capacity has become a hurdle in most recent generation of mobile devices. As the average power consumption of the devices increases dramatically, there is still no breakthrough with battery capacity of mobile devices itself, leading to shorter and less-predictable battery life. According to a user study [49], the most wanted feature in the future mobile device is *days of battery life during active use*. A brief quantitative as well as historical account of the increase in battery capacity with reference to total phone weight is presented in [50]. A more detailed view of the advances in this direction can be found in [51, 52, 53]. In brief, it is imperative to develop techniques to reduce battery energy consumption for MSs with heterogeneous connections. This is indeed the motivation for our work in Paper B and Paper D.

3.2 Access Mechanisms for IEEE 802.11-based WLANs

The major goals of the MAC mechanism in IEEE 802.11 WLANs are the provision of reliable data delivery service to end users, fair access to the shared wireless medium, and to protect the data packet that it is under delivering. There are two functions in IEEE 802.11 WLANs and they are briefly summarized below since they constitute the basis for our analysis in Paper B.

3.2.1 Distributed Coordination Function

Distributed Coordination Function (DCF) [54] is the fundamental MAC mechanism, mandatory for 802.11 WLANs, whose underlying protocol is Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA). Fig. 3.1 illustrates the concept of CSMA/ CA.

The basic CSMA/CA algorithm works as follows. A station listens to the radio medium and checks the status of the channel when it has packet to send. If the channel is busy, the transmission will be deferred. If the channel has been idle for Distributed Inter-Frame Spacing (DIFS), a station can transmit after a backoff period. The backoff period is randomly selected from a uniformly distributed contention window. An explicit ACK (i.e., acknowledgment) frame is sent by the receiver upon the successful reception of a unicast packet. If the ACK is not received within the ACKTimeout interval, the sender shall retransmit the packet, up to the RetryLimit (an integer number which is 4 for short RetryLimit and 7 for long RetryLimit).

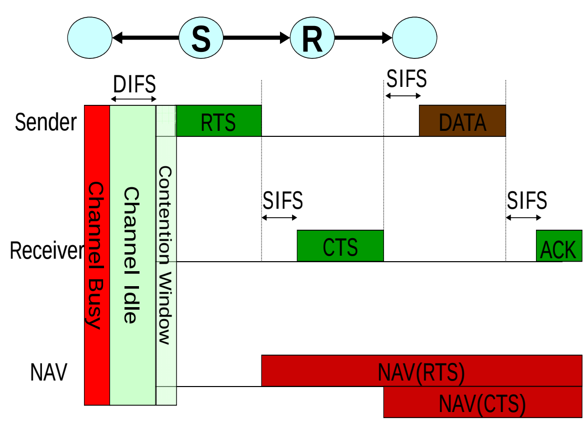


Figure 3.1: CSMA/CA mechanism illustration.

To avoid collision, the other stations stay silent for a Network Allocation Vector (NAV) period when there is an ongoing frame transmission. Moreover, the binary exponential backoff intervals are used to reduce collision probability. As the basic two-way handshake mechanism, CSMA/CA+ACK is used to ensure the reliability of transmissions. Additionally, a four-way handshake mechanism, CSMA/CA+RTS/CTS, where RTS and CTS stand for Request To Send and Clear To Send respectively, is used to eliminate the hidden terminal problem. In the RTS/CTS enabled mechanism, a sender first transmits an RTS frame. If a corresponding CTS frame is received which means that the receiver is ready, the DATA frame will be transmitted, followed by an ACK frame from the receiver.

The success of DCF is credited to its simplicity and distributed nature which provides asynchronous data delivery service. However, the major defect of DCF is its inefficiency, which becomes even worse when the number of contending stations increases. The performance degradation is ascribed to severe collisions and bandwidth waste on idle backoff slots.

3.2.2 Point Coordination Function

Point Coordination Function (PCF) [54] is implemented on top of the DCF and it allows an AP to poll other stations where the AP is the polling master. The AP sends *beacon* frames periodically and the frames are separated by Short Inter-Frame Spacing (SIFS) which is smaller than DIFS. No collisions occur during the Contention Free Period (CFP) because during this period stations can transmit only when they are polled by the AP. All unicast transmissions must be acknowledged. Meanwhile, the AP can also tell other stations to sleep to save battery lifetime and holds temporarily frames destined for the sleeping stations.

Moreover, PCF is an optional access mechanism but it utilizes wireless channel much more efficiently, especially when traffic load is heavy. It has been shown that the PCF achieves higher saturation throughput than the DCF does. However, due to the null-polling overhead, the PCF raises more access delay when stations have rare packets to send. With several open problems unsolved, the PCF protocol has not been widely deployed in real-life [55].

Comparatively speaking, DCF provides better latency performance under a light-load condition, while the PCF can achieve higher system capacity [56]. A thorough description of the the DCF-based Contention Period (CP) and PCF-based CFP can be also found in [57].

3.3 Energy-efficient Mobile Networking: A Retrospect

From the connectivity perspective, one can observe that the current Internet paradigm is based on the idea of placing intelligence at the hosts, which then need to have end-to-end always-on connections and conversational mechanisms to exchange information [58]. [50] explains why the host-centric paradigm introduces intrinsic energy inefficiency and argues that patches to make currently deployed protocols energy-aware cannot provide an order of magnitude increase in energy efficiency.

Furthermore, extending the operational time of mobile devices has proven to be a difficult task and may not be solved without a step-change in mobile networking. An often cited power budget for a 3G mobile phone streaming a 384 kbps video includes 1.2 W power draw for the network interface/modem, 1 W for A/V and backlight display, and 0.8 W for Central Processing Unit (CPU) and memory operations [50], [59]. This indicates that the power saving advances need to be achieved on *several fronts*, i.e., simply optimizing one component cannot lead to significant reduction. Moreover, we need to explore further methods from other perspectives rather than the component-level power reduction in order to reduce the energy consumption of mobile devices. Paper B makes an effort towards this direction.

In the rest of this section, we first summarize the efforts on component-oriented mobile battery lifetime extension and then briefly present related work with respect to mobile-oriented energy consumption for heterogeneous connections.

3.3.1 Component-oriented Efforts for Mobile Battery Lifetime Extension

Recently, several approaches have been proposed and implemented to reduce communication device power consumption, and thereby extend battery lifetime [60]. A few of these are discussed below.

Efforts towards optimizing WLAN communication device power consumption mainly utilize the Power Saving Mode (PSM) for Network Interface Card (NIC) power management [54]. In [61], the authors present a smart PSM for IEEE 802.11 WLANs. Another improvement is made to make energy efficient Voice over Internet Protocol (VoIP) traffic over WLAN based on the PSM [62].

Moreover, the authors in [63] propose a receiver side coalescing scheme allowing Ethernet NICs to perform deep packet inspection to improve TCP/IP processing efficiency. However, their scheme does not regulate interrupts, thus the platform wake up pattern remains the same.

There are also several studies which propose to opportunistically put both NICs and the platform to sleep mode, when the platform is inactive. As the platform consumes significantly less amount of power when in the sleep state, this allows a further power saving. For example, [64] allows the platform to go into the sleep mode when it is completely idle, and it wakes up when there are packets waiting at the BS. A similar philosophy, however, with more advanced packet filtering technology is proposed in [65] to ensure connectivity even when the host is in the sleep mode. The main drawback of such approaches is that they save power only when the platform is idle, i.e., has no active load.

As Internet traffic is bursty in nature, [66] observes that by taking the advantage of traffic regulation, platform wake-up events can be reduced, thereby reducing power consumption. To minimize power consumption by designing energy efficient platform components, [67] and [68] present proposals from the technological perspective at the digital circuit level.

3.3.2 Mobile-oriented Energy Consumption for Heterogeneous Connections

For research efforts in HNs, attention has also been paid on the development of handover policies. So far, little work has been done on the MS's energy consumption based on the amount of traffic in a selected network within an HN.

The IEEE 802.21 standard [69] supports algorithms to enable seamless handover across different networking technologies. This is also known as Media Independent Handover (MIH) or vertical handover. The procedure on how to perform handover to and from IEEE 802.3, IEEE 802.11, IEEE 802.15, IEEE 802.16, 3rd Generation Partnership Project (3GPP) and 3GPP2 networks through different handover mechanisms is provided in the standard. Many vendors have participated in the development of the IEEE 802.21 standard and built wireless products.

The authors in [70] investigate handover performance between 3G mobile systems and WLAN access networks through simulations. The WLAN signal level thresholds are used as the handover criteria by assuming that the MSs support the IEEE 802.21 cross layer architecture. Given that the economic cost and data rate are the major factors in their consideration, the authors of [70] are in favor of using WLAN over UMTS. However, the energy consumption of the MS is not considered as a decision-making factor for handover between these networks. Furthermore, an experimental testbed is developed by the authors in [71] to implement certain parts of an IEEE 802.21-based framework. They report that the IEEE 802.21-assisted handover performs reliable and proactive handover by tolerating longer delays and higher packet loss to an acceptable level. The effect of energy consumption on the handover decision is, however, not investigated in their work. Moreover, a relatively experimental study focusing only on the application layer of the communication protocol is presented in [72]. A handoff algorithm based on energy consumption measurements of UMTS and IEEE 802.11 WLAN networks is proposed. However, the effects of distance between the involved transceivers are not considered on the battery energy consumption.

Furthermore, energy-efficient network activities in mobile phones supporting multiple wireless technologies are investigated in several studies. For example, [73] develops strategies to intelligently switch between WiFi and bluetooth. An architecture is proposed in [65] which uses the GSM radio to wake up the WiFi radio upon an incoming VoIP call to leverage the battery quality and energy efficiency of WiFi while keeping its scanning cost low.

Another study [74] demonstrates that intelligently switching between WiFi and GSM substantially reduces energy consumption. An algorithm is designed by the authors to predict WiFi availability, by letting the device scan WiFi APs only in areas where WiFi is available with high probability. In this way, unnecessary scanning in case of poor WiFi availability is avoided, thus the algorithm is more energy consumption friendly.

To summarize, neither the IEEE 802.21 standard itself nor existing studies suggest any algorithms or schemes to base handover on the battery energy information of involved device(s). In Paper B, we propose link distance based energy consumption methodology for MS energy consumption calculation, which constitutes the basis for designing energy consumption based handover algorithms.

3.4 Analysis of Link Distance-based Mobile-centric Energy Consumption: The Proposed Model

Although the related work mentioned above covers many aspects of energy consumption of MSs from different perspectives, no existing work has addressed the question on whether there will be any advantage or disadvantage to select different connections for uplink and downlink traffic within an HN scenario. In Paper B, we focus on this issue and suggest that selecting different connections for different traffic directions may save energy for the involved MS.

3.4.1 Hybrid Network Scenario and Possible Connections

The network considered in Paper B is a wireless HN where the MSs have heterogeneous connections. The envisaged HN scenario is re-drawn as Fig. 3.2. The MS may access the Internet through one of the three alternative links, i.e., a direct Low Data Rate (LDR) cellular link, a direct High Data Rate (HDR) WLAN link through an AP, and a combined HDR link via a Relay Station (RS) towards the BS. The link between the MS and the RS is part of an ad hoc network.

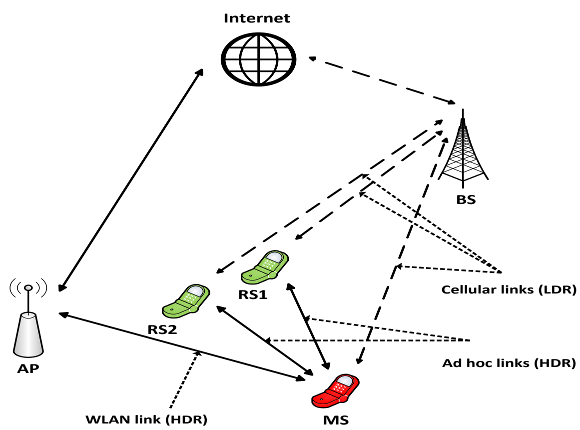


Figure 3.2: A hybrid network scenario with three heterogeneous links.

In Fig. 3.2, there are various alternative connections for the MS to establish communication with the Internet for both uplink and downlink traffic. The connections considered in Paper B are:

- direct *cellular connection* in which the MS directly reaches the BS for both uplink and downlink,
- *WLAN connection* in which the MS uses the AP to reach the Internet for both uplink and downlink,
- Uplink Relayed Downlink Relayed *(URDR) connection* in which the MS uses one-hop RS to reach the BS for both uplink and downlink,
- Uplink Cellular Downlink WLAN *(UCDW) connection* in which the MS uses cellular link for uploading and WLAN link for downloading,
- Uplink WLAN Downlink Cellular *(UWDC) connection* in which the traffic direction is opposite to that used in the UCDW connection,
- Uplink Relayed Downlink WLAN *(URDW) connection* in which the MS uses an RS to reach the BS for the uplink and uses AP for the downlink, and

- Uplink WLAN Downlink Relayed *(UWDR) connection* in which the traffic direction is opposite to that used in the URDW connection.

The AP and the BS in the scenario are assumed to be without energy constraint while the MS and the RSs are battery powered and hence have limited energy.

In what follows, we summarize the underlying methodologies used in Paper B for the analysis of MS energy consumption over individual links. Furthermore, the major contributions of Paper B are also briefly presented. More details about our analysis are given in Paper B.

3.4.2 Power Consumption in the LDR Cellular Link

For the power consumption calculation in the cellular link, one or more MSs are considered to be within the coverage of a BS with CDMA connectivity. The power consumed per bit by the MS, P_c, is divided into two parts, i.e., the electronic power, P_{et}, needed to radiate P_t amount of transmission power from the antenna and the electronic processing power, P_{er}, consumed for receiving and decoding the received signal [75]. As more circuitry is needed to receive and decode the signal [76], the consumed power by the MS battery can be expressed as

$$P_c = k_1(P_{et} + P_t) + k_2(2.5P_{er}), \tag{1}$$

where k_1 and k_2 are mutually exclusive binary digits. The transmission power of an MS is obtained from the simplified path-loss model, i.e., $P_r = AP_t/d^{\alpha}$, where P_t and P_r are the transmitted and the received power respectively, α is the path-loss coefficient, d is the distance between the transceivers, and A is a unitless constant dependent on the involved antenna characteristics. The energy consumed per byte, E_c, can then be calculated by using $E_c = 8 \times P_c/R_c$ where R_c is the supported data rate.

3.4.3 Power Consumption in the HDR Ad Hoc Link

The MS power consumption for the hybrid link (i.e., MS-RS-BS, and vice versa) can be obtained by adding the individual power consumption in the cellular link and the ad hoc link. This is due to the fact that a station cannot transmit and receive simultaneously.

In the ad hoc network part of the overall HN, all stations are assumed to access the channel through the DCF [57]. However, since the MSs in the ad hoc network can use the AP as well, the overall access control is regulated by the AP through

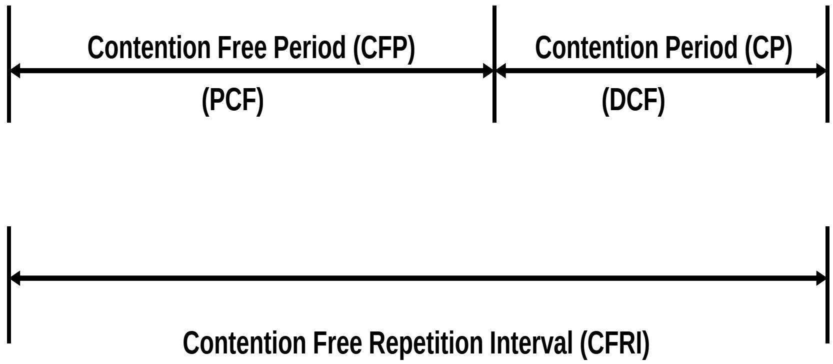


Figure 3.3: Illustration of contention free repetition interval.

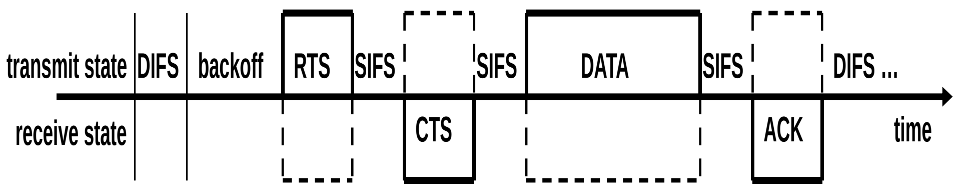


Figure 3.4: Interaction between transmitter and receiver during a CP.

PCF, which runs on top of the DCF. The AP dictates the duration of the CP and the CFP, which collectively constitute the duration of Contention Free Repetition Interval (CFRI). DCF is used to access the channel during the CP. Fig. 3.3 illustrates the time relationship between DCF and PCF in a CFRI.

For our calculation of MS power consumption during its transmission and reception states, Fig. 3.4 is used as the benchmark. The corresponding energy consumption values are obtained by multiplying the calculated power in a state with the respective time spent in that state.

3.4.4 Power Consumption in the HDR WLAN Link

Fig. 3.5 is used to calculate the MS power consumption during the CFP. A *beacon* frame announces the start of the CFP. For a detailed description of the individual frames mentioned in Fig. 3.5, the reader is referred to [57] and [54].

Once the power consumption expressions of the MS in the individual links are obtained, the power consumption for the mixed connections can be obtained collectively.

The results of Paper B suggest that MS energy consumption can be reduced by considering distinct links for uplink and downlink traffic. To increase throughput while maintaining minimal power consumption, factors such as the remaining bat-

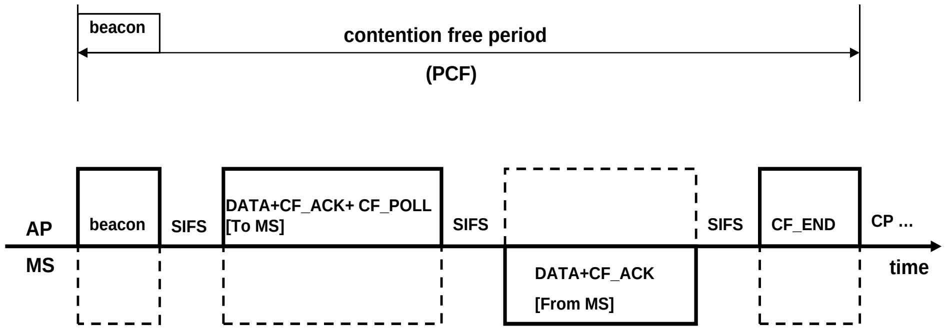


Figure 3.5: Interaction between transmitter and receiver during CFP.

tery energy of the mobile station and the amount of bytes communicated over a link should be taken into consideration for selecting these links.

3.4.5 Major Contributions of Paper B

Although the analytical model proposed in Paper B is simplistic, it gives an initial foundation for power consumption analysis in hybrid networks with heterogeneous connections. Furthermore, the idea of separate power consumption calculations for uplink and downlink in order to obtain an insight on the potential benefits of using different links for data transfer is original. This is a novel concept, and it is not found in the existing literature, to the best of our knowledge. It is pertinent to mention that in-depth investigation is nevertheless encouraged towards this direction.

This paper also provides insight on *mobile-centric* energy-aware link selection. Furthermore, the paper can be used for developing handover strategies, especially beneficial from MS's perspective. A few possibilities for handover decision are discussed in Paper B. However, more comprehensive handover strategies need to be developed by taking into consideration multiple optimization parameters in order to minimize the MS energy consumption in an HN. These parameters may include the distance between transceivers in a certain link, the channel conditions for individual links, the speed of data transfer for a certain link, etc.

Another important contribution of the paper is that it initiates the potential of *energy-aware* mobile-centric handover decision. The energy consumption of the device in a certain link, before and after handover, is not yet considered as a handover parameter by the IEEE 802.21 standard. The manufacturers of the devices based on the IEEE 802.21 standard will be interested in developing algorithms which could minimize device energy consumption, and herein this paper serves as an initial guideline.

3.5 Chapter Summary

Battery energy conservation is an imperative issue due to the mismatch between battery technology development and the explosion of data-centric mobile applications. In this chapter, we have discussed several issues concerning the battery energy conservation for MSs operating in HNs. Related efforts, made by both industry and academia, are summarized. As WLANs (in both infrastructure mode and ad hoc mode) and cellular networks are the most popular constituents of HNs, the energy consumption of an MS operated in these networks is analyzed. A brief summary of the access mechanisms used in WLANs is also presented since this is essential for our analysis in Paper B.

A scenario that an MS operates in an HN consisting of a cellular, an infrastructure-based WLAN, and an ad hoc network is investigated in-depth. Correspondingly, the baseline for energy consumption analysis for MSs operated in such an HN is outlined for each link. This constitutes the basis for the work carried out in Paper B. The chapter concludes with a brief discussion on the major contributions as well as the impact of Paper B.

Chapter 4

Towards Green Communications in Cellular Networks

4.1 Energy Consumption in Cellular Networks: An Overview

This section gives an overview on energy consumption in cellular networks. Generally speaking, energy consumption of cellular networks can be viewed from two different but related perspectives, i.e., the operator's (or system-level) perspective and the MS's perspective.

4.1.1 Energy Consumption from Operators' Perspective

As electricity bill is approximately 20 percent of the total operational expenditures for mobile networks, operators foresee energy efficiency of their networks as an important component to reinforce their business competitiveness.

In a cellular network, BSs are the most energy demanding component, responsible for the consumption of about $50 \sim 70$ percent of the total network energy. Therefore, there is a great potential to save a reasonable amount of network power if the energy consumption of BSs could be reduced. A direct way to increase energy efficiency of cellular networks is through network deployment using environment-friendly topologies [77, 78, 79]. For instance, sufficiently dense network deployment, consisting of small low-power BSs, is expected to reduce the overall network energy consumption in contrast with the deployments of few high-power BSs [77]. As network coverage must be ensured at any time in all service areas, hierarchical cell structures, where small cells such as micro- and pico-cells are used for capacity enhancement, have been shown to be promising solutions [78], [79]. This type of technologies can increase energy efficiency, however, at a cost of deployment expenditure.

Moreover, due to factors like user mobility and behaviour, traffic in cellular networks experiences significant fluctuation in space and time. Traffic load is usually higher during daytime in office districts and during night time in residential areas. Hence, there are always some BSs operating under low load, making static cell deployment a less efficient solution. Such fluctuations in traffic can be a serious issue for future generation micro-, pico-, and femto-cells. Therefore, at the management level, a network-level BS power management is required where many BSs cooperate in an intelligent manner to bring down the overall average BS power consumption. [80] gives a methodology to dynamically adjust the cell size in a multi-layer cellular network architecture.

Power Amplifier (PA) is another major source for energy consumption in a cellular communication system. The energy efficiency of a PA is calculated by the ratio between the AC power input and the generated RF output power. The radio part of a BS consumes about 80 percent of energy, out of which PA consumes about 50 percent [81]. Depending on the system used (e.g., UMTS, GSM, CDMA), the efficiency of a PA can be between $5 \sim 20$ percent. However, higher efficiency PAs have now been reported [82, 83], with efficiency reaching over 50 percent. Additional efficiency can be achieved by shifting to the switch-mode PAs rather than traditional RF amplifiers. Switch-mode PAs generate very little power as heat, resulting in a highly efficient power supply. The total component efficiency of such devices can reach around 70 percent [81].

4.1.2 Energy Consumption from MSs' Perspective

In order to achieve true green communication in networks, energy conservation for MSs also needs to be considered. While mobile devices in cellular networks always need to be associated with their base stations, WiFi incurs a high initial cost of associating with an AP. However, as WiFi, e.g., as an MS interface, typically uses the power-save mode, the cost of maintaining this association is low. When associated, the energy consumed by a data transfer is proportional to the size of the data transferred and the transmission power level. The same observation applies also to ad hoc networks.

Moreover, there are further energy consumption components in a mobile device which are dependent on the type of the platform used. These components include the CPU, memory, chipset, screen etc. However, their energy consumption can be taken collectively as the energy consumed for electronic processing. In addition, other factors also indirectly affect the battery energy of an MS, including the type of application(s) running, the ratio between the amount of uploaded and downloaded data, etc.

Furthermore, many approaches aimed at reducing energy consumption in cellular networks rely on the status of BSs going to sleep mode. However, this may cause the MSs in the *affected* area to transmit at a higher power level in order to reach the active BSs. Hence, green communication from the BS's point of view may not be seen as green from the MS's point of view. Although many proposals aimed at reducing MS power consumption exist in the literature [84] [85], there is little reported work on joint optimization of BS and MS energy consumption. For example, [86] shows how optimizing BS transmission power alone in downlink can lead to increased MS power consumption. The authors propose sub-optimal algorithms to find the right balance in achieving energy savings in both the MS circuit power consumption and the BS transmission power consumption by exploiting the tradeoff between energy and delay. Another proposal [87] presents a new architecture for wireless networks, aiming at minimal emission from MSs. The authors suggest to equip each transceiver BS with extra specially dubbed *green antenna*. The MSs near the green antenna can transmit at lower power, thereby reducing their own energy consumption and generating less interference to other users. The green antennas do not produce additional radiation, since only uplink trafic is relayed.

Based on the literature survey in the above two subsections, we believe that joint consideration of BSs' and MSs' energy consumption in cellular networks would achieve optimal solution for green cellular communications. Anyhow, the research work performed by Paper C is more towards the operator's perspective. We have however considered MS traffic load in our scheme to be presented later in this chapter.

4.2 Energy Reduction Approaches in Cellular Networks

Having summarized energy consumption in cellular networks and MSs, this section briefly presents the state-of-the-art techniques for energy conservation from architectural, emerging technologies, and network planning points of view.

4.2.1 Architecture-level Energy Saving in Cellular Networks

In order to achieve system energy minimization, Discontinuous Transmission (DTX) and cell zooming are two typical feasible techniques [10] [15]. Furthermore, as PA dominates the energy consumption of a BS, the energy efficiency of PA needs to be addressed in order to improve the hardware design of BSs.

Base stations and MSs in the current cellular network architecture are required to periodically transmit pilot signals. An intuitive way to save BS energy is to switch off the transceivers when there is no need to transmit or receive. The LTE standard utilizes this concept by introducing power saving protocols like Discontinuous Reception (DRX) and DTX. As continuous transmission and reception consume a significant amount of power, DRX and DTX are an attractive way for BS power consumption reduction. Anyhow, such power saving protocols for BSs are still in the design phase and have not been deployed in current mobile systems.

Moreover, a concept of Self-Organizing Networks (SONs) is introduced in 3GPP standard to add network management to improve network performance and flexibility as well as to reduce cost [88]. Different use cases of SON, e.g., cell outage management, load balancing etc, are discussed in [89]. [15] and [16] are among the initial ones to explore the performance of self-organizing techniques.

In [90], the authors propose Intelligent Cell BrEathing (ICE) to optimize the utilization of green energy (through energy harvesting [91]) in future cellular networks, and therefore minimizing the energy consumption from the main grid in order to reduce the operator's cost.

Furthermore, a distributed algorithm where the BSs can take turns in reducing their power through the exchange of information about their current power levels is proposed in [92]. Moreover, the authors of [93] and [94] propose an association among the powered-on and powered-off BSs to rearrange the energy configuration using the concept of energy partition.

From the implementation point of view, [10] presents a framework for cell zooming which includes a cell-zooming server to sense the network traffic and channel quality etc. Such cell-zooming servers can be distributed in BSs. The authors in [80] propose dynamic self-organization of the cellular layers using timed sleep mode, user location prediction, and reverse channel sensing.

Adaptive transmission power is another straightforward method for energy saving [95]. The effect of cell size on energy saving and system capacity is analyzed in [96]. The authors demonstrate the effectiveness of small-cell based mobile communication systems in terms of energy efficiency. However, the benefit of energy saving for operators is generally not significant with this technique because of the comparatively huge energy consumption of the involved BSs.

4.2.2 Energy Conservation through Emerging Technologies

The recent developments in the area of cooperative communication have the potential of improving energy efficiency through distributed signal processing. In the following, we discuss some of these technologies used in cellular networks.

Covering distant users via direct transmission can become expensive in cellular networks due to high transmission power requirements caused by heavy path losses and shadowing etc. Cooperative techniques are able to combat channel fading by covering coverage holes [97]. Indeed multi-hop communication between MS and BS through shorter links [98] reduces transmission power requirement, and hence can also extend MS battery lifetime [99]. Furthermore, using multi-hop in CDMA cellular networks has been shown to reduce average energy consumption per cell [100].

In real-life, green communications using cooperative techniques can be achieved in two ways: 1) through the installation of fixed relays within the network coverage area, and 2) through the exploitation of existing users to act as relays. Chapter 2 also sheds light on the concept of relaying.

Fixed relays in a cellular network provide a feasible solution for network power reduction. For an Additive White Gaussian Noise (AWGN) channel with a path-loss coefficient of 4, [101] shows that by increasing the number of BSs by a factor of 1.5 in a unit area, we can reduce the transmission power by a factor of 5, still keeping the same SNR level. This means higher density of BSs with less energy consumption. As installing new BSs to have higher BS density can be expensive, deploying relays instead is more economical and introduces less complexity to the network. On the other hand, relays can be connected to the BSs wirelessly (unlike BSs which need to be connected to the backhaul in a wired way). The authors in [101] discuss the possibility of green communication in a cellular network using fixed relays.

User cooperation was first introduced in [102] and has been shown to be able to increase data rate. However, whether user cooperation can contribute towards achieving green communication in a mobile network is still an open question. The reason is simple: increased data rate for one user comes at the cost of energy consumed for another user acting as the relay. Therefore, if no incentives exist, limited battery energy of MSs in a mobile network renders this option unattractive, in general. In [103], the authors use a game theory-based approach to give incentive to the MSs to act as relays when they are idle. They also pose the question of whether or not the user cooperation is advantageous from the energy efficiency point of view.

4.2.3 Network Planning for Energy Conservation

In cellular networks, smaller cells such as micro-, pico-, and femto-cells are more power efficient due to their sizes. [82] reports that with only 20 percent of customers covered by pico-cells, the joint deployment of macro- and pico-cells in a network can reduce the energy consumption of the network by about 60 percent. However, a careful deployment strategy needs to be investigated before any actual deployment of such networks. Deploying too many smaller cells within a macro-cell may cause the macro-cell BS to operate under low load conditions, thereby wasting power. Moreover, the impact of different deployment strategies on power consumption of cellular networks is investigated in [104]. Their simulation results suggest that under full traffic load the use of micro BSs has a moderate effect on the area power consumption of a cellular network and it strongly depends on the offset power consumption of both macro- and micro-BSs. In [105], the authors investigate possible scenarios for joint deployment of macro- and femto-cells.

Furthermore, the authors in [106] present an approach to create links between fully centralized cellular and decentralized ad hoc networks in order to achieve more efficient network deployment. Their goal is towards self-organizing small-cell networks. In [107], the authors show that depending on the voice traffic model, the large-scale femto-cell deployment can provide an average power saving of 37.5 percent by completely turning off femto-cell BSs' transmission and processing when they are not involved in an active call.

As another alternative, BS sleep mode design has gained popularity in recent years [108], [19]. One of the earliest papers to advocate for energy conservation in networks is [109]. The authors suggest to put network components to sleep in order to save energy. The idea of BS sleep mode design in both the time and spatial domains according to LTE standards is proposed in [108]. The authors show that by switching off the BS *transmission* during certain sub-frames in each radio frame, up to 90 percent RF energy reduction can be obtained because of less control overhead for signaling. Another interesting energy consumption minimization approach is presented in [110] where a number of resources in a system can be shut down for given traffic scenarios in either a semi-static way or a dynamic way. The authors show that greater potential of energy saving exists using dynamic scheme. According to these authors, the major issues in sleep mode-based methodologies are (i) the activation time that may result in blocking of new or handover calls and (ii) the ping-pong effect which results in unnecessary ON/OFF oscillations.

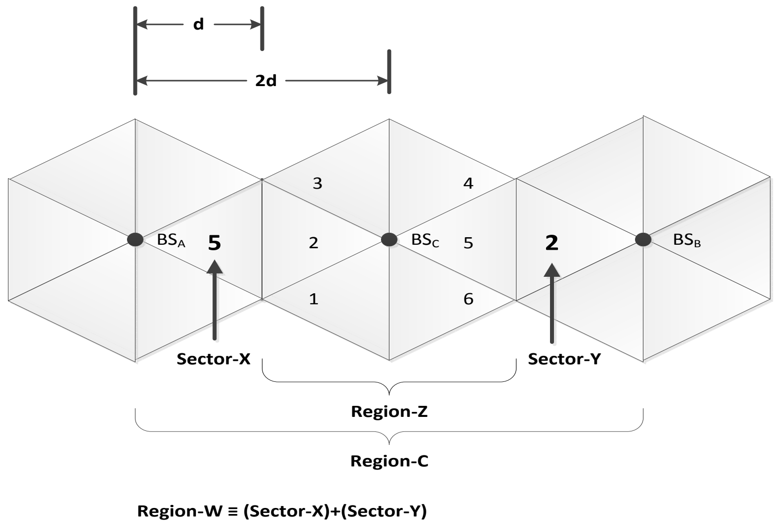


Figure 4.1: Network scenario used in Paper C, where the BS in Region-Z can be switched off under light traffic conditions.

Based on the above literature survey, we observe that although some BS switch-off (or sleep-mode) approaches exist [16] [111] [10] [15], they mainly take 24 hour traffic scenario as the baseline for their mode shift. These approaches do not consider the fact that traffic is not hour-based and may vary over short time. In Paper C, we propose two approaches where the switch-decisions rely only on traffic load instead of the time of a day. The ping-pong effect is also avoided in our schemes. In the following two sections, we summarize the main idea of these two solutions.

4.3 A Deterministic Approach for Power Saving

As discussed above, the approach based on switching the BSs off (or taking them to sleep mode) under low traffic load conditions is a promising alternative. Paper C proposes one such approach and tackles the problem from two different angles, i.e., either in a deterministic fashion or in a probabilistic fashion. In this section, we discuss only the deterministic approach, and the probabilistic method is left to the next section.

The deterministic approach proposes to let the middle cell go to sleep mode under light traffic conditions and correspondingly to cover its users by increasing the transmission power of an antenna from each of the two neighbouring cells, under low traffic load conditions. The network scenario is depicted in Fig. 4.1, where a network consisting of three micro-cells and corresponding BSs is considered. Each of the cells is divided into six sectors and the users in each of the sectors are covered

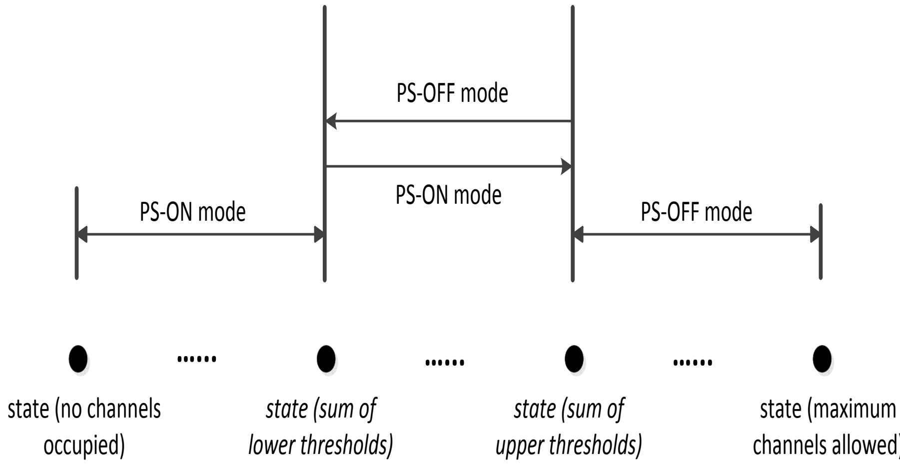


Figure 4.2: Power saving policy used in Paper C, where a hysteresis region is introduced to avoid the ping-pong effect.

by a dedicated sectored antenna. The BS for the middle cell, BS_C, also referred to as Region-Z, can be switched off under low traffic conditions. The users of this cell are then covered by increasing the transmission power of the antennas (from d to $2d$) for Sectors 5 and 2 (also referred to as Sector-X and Sector-Y, respectively, in the figure) of BS_A and BS_B, respectively. The total region, consisting of Region-Z and Sectors X and Y is collectively referred to as Region-C in the figure.

The maximum number of allowed channel occupancy in an individual sector is r. As two sectors are used to cover the switched-off cell, the total allowed channel occupancy for these two sectors plus six original sectors from the middle cell, becomes $2r$. The traffic in the network is analyzed using Markov Chains (MCs) by dividing the network into a few regions and modeling the traffic in each region as a birth-death process, assuming Poisson arrival rate and exponential service times.

There are two operation modes defined in the system, i.e., PS-ON and PS-OFF. The PS-ON mode means that the system shuts off the middle cell and saves power (of the BS from the middle cell) while the PS-OFF mode indicates the normal operation of all BSs in the system. The deterministic power saving scheme proposed in Paper C is based on sojourn time analysis. In this scheme, once the system reaches a triggering state, the mode is *deterministically* changed, hence the name *deterministic*. A power saving policy is developed based on the channel occupancy threshold in the network.

Fig. 4.2 illustrates the power saving policy proposed in Paper C. It is worth mentioning that the states mentioned in Fig. 4.2 indicate the collectively allowed channel occupancy in Region-C. As shown in the figure, the total allowed channel occupancy for the covered sectors, $2r$, is divided into three groups of states by two thresholds. The system remains in the PS-ON or the PS-OFF mode if the number of channels occupied is between 0 and the lower threshold or between the upper threshold and $2r$, respectively. The hysteresis region in the middle part of the figure, where the system keeps its current mode according to a few rules, is designed to avoid the ping-pong effect and is described in details in Paper C.

The power consumption of the network in each mode is calculated using the mean time spent in the respective states of each mode which is obtained through a matrix whose entries represent the expected time periods the MC (representing the traffic in the covered sectors) stays in state j given that it starts from state i.

The effects of blocking probability (represented with relation to the Grade of Service (GoS)) and offered traffic load, λ/μ, on the network total transmission power consumption are discussed in Paper C. Furthermore, the effects of hysteresis region boundary shifting are also explored in the paper. The numerical results demonstrate that we can achieve a considerable amount of network power saving using the proposed scheme.

It is also worth mentioning that some of the results not shown in Paper C are contained in [112]. As an example, we illustrate such a result on the effects of hysteresis region length variation on power consumption in Fig. 4.3. The figure illustrates the total transmission power consumption in the PS-ON and the PS-OFF modes for $\lambda/\mu = 0.5$ with different upper and lower thresholds for channel occupancy, as $t_l + t'_l$ and $t_u + t'_u$ respectively, where the total number of channels is considered as 17.

To study the effects of hysteresis region size, the length of the hysteresis region [i.e., $(t_u + t'_u) - (t_l + t'_l)$] is reduced to 5 in Fig. 4.3(b) from being 6 in Fig. 4.3(a). We observe that the difference in respective transmission power is also lower for smaller hysteresis region. This is because that as the hysteresis region is shrunk as well as $t_l + t'_l$ and $t_u + t'_u$ are brought closer to the middle state value, the system spends roughly equal fraction of time in both the PS-ON and the PS-OFF modes. Furthermore, due to a lower value of $t_u + t'_u$, there are more number of states within the pure PS-OFF region (since now $t_u + t'_u = 10$). This indicates a relatively higher fraction of time spent in the PS-OFF mode, resulting in a lower amount of transmission power consumption in Fig. 4.3(b) than in Fig. 4.3(a). For a detailed description of the other results, the reader is referred to [112].

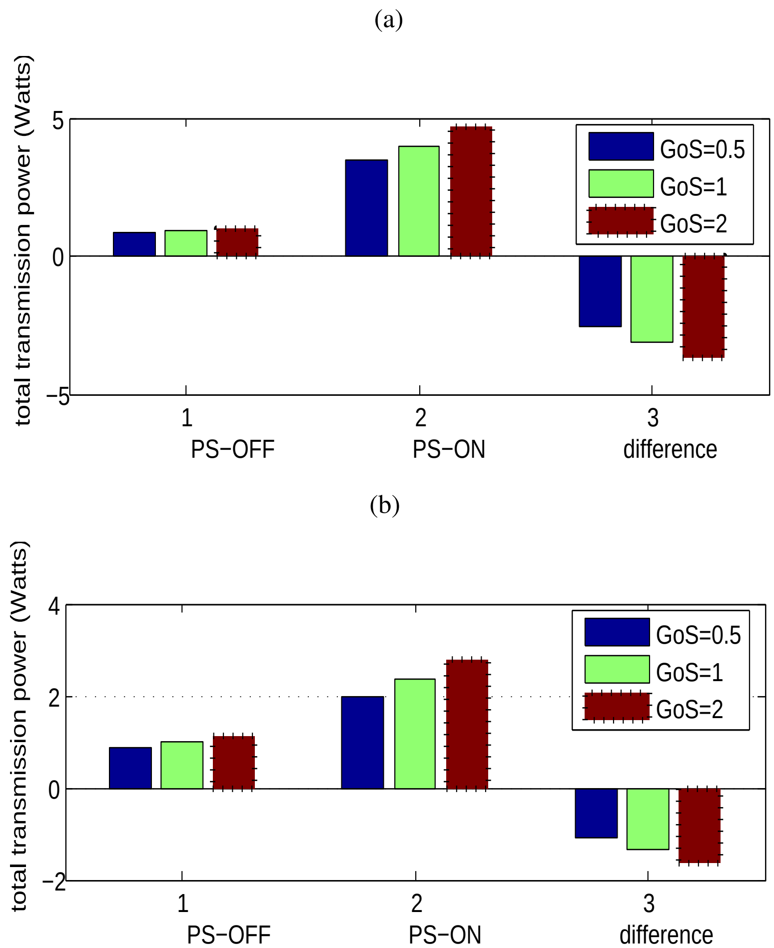


Figure 4.3: Total transmission consumptions: (a) $\lambda/\mu = 0.5$, $t_l + t_l' = 7$, $t_u + t_u' = 12$, (b) $\lambda/\mu = 0.5$, $t_l + t_l' = 6$, $t_u + t_u' = 10$.

4.4 A Probabilistic Approach for Power Saving

This section outlines the probabilistic approach proposed in Paper C for network transmission energy consumption analysis. The optimization of long-term network transmission power consumption cost is also of operators' interest. In order to achieve this, we may utilize reinforcement learning technique and therefore a probabilistic approach is proposed. By probabilistic, we mean that *the decisions to switch to the PS-ON or the PS-OFF mode are taken with certain probability*, instead of shifting to a mode deterministically.

For the analysis of this approach, we use an FMDP to optimize the transmission power consumption of the same given three-cell network scenario. An FMDP is basically a reinforcement learning methodology which satisfies the Markov property [113]. It is defined by the state and action spaces and by one-step dynamic of the environment. All the states that the process can assume constitute the state space while all the possible actions that can be taken from each of these states constitute the action space of the FMDP. A transition to another state with certain probability may (or may not) be triggered by taking an action from the action space. In addition to that, for any current state and action pair, there is an expected value of associated cost.

A state in the process is one of the ranges of channel occupancy, as defined in Paper C. The action space consists of three actions, i.e., 1) to switch to the PS-ON mode, 2) to switch to the PS-OFF mode, or 3) to keep the current power mode unchanged. These actions are taken with certain probabilities to optimize the expected cost which is defined as the transmission power consumed in a state. For this purpose, an initial policy is defined, which converges to the optimum policy through iterations. A linear program is used for minimizing the cost function.

The results for this part of the paper are obtained for the optimized power consumption of the system in both the PS-ON and the PS-OFF modes, but for transmission power only. The results demonstrate that a lower limit on the long-term network transmission power can be obtained using the proposed FMDP-based analysis. Finally, a comparison between the deterministic and probabilistic approaches is also given in Paper C.

4.5 Distinguishing Aspects of Paper C

As hinted in Secs. 4.1 and 4.2, a few studies in the literature consider BS switch-off methodology in macro-cell scenarios. However, this type of approaches may not be reasonable because the transmission radius of a macro-cell BS is large, thereby infeasible to further extend the transmission radius to cover the switched-off cell. To avoid this problem, our envisaged scenarios target only at micro- and pico-cells, for which transmission power increase may not be a problem due to smaller transmission radii.

Furthermore in our scheme a hysteresis region based on the channel occupancy thresholds of the defined regions is defined. This enables us to diminish the influence of the ping-pong effect. Other approaches in the literature generally ignore this aspect while proposing BS switch off schemes.

Another distinct feature of the proposed schemes in Paper C is that our methodologies are applicable to all times of a day, instead of merely considering the peak traffic hours which is the case in quite a few related papers. Furthermore, a novel technique based on FMDP analysis is proposed for the calculation of minimum network transmission power. This approach gives the operators an insight on the minimum long-term power consumed in their networks.

4.6 Chapter Summary

In this chapter, we discussed issues relevant to cellular network energy conservation from a system point of view. Related efforts on these issues, made by both industry and academia, were briefly presented. In Sec. 4.1, a broader view on energy consumption in cellular systems was presented from both the operator's and the MS's perspectives. The state-of-the-art energy reduction techniques for cellular networks from architectural, enabling technologies, and network planning points of view were given in Sec. 4.2.

Considering that BSs are the most important components for network energy consumption, and traffic intensity varies over time, this chapter presents two approaches targeting at switching off BS under light traffic load conditions, for the purpose of network energy saving. More details about the approaches, the analytical work as well as the numerical results are given in Paper C.

Chapter 5

Markov Decision Process for Battery Energy Optimization

5.1 Introduction to Mobile Battery Consumption Estimation

Another aspect of green communication deals with energy consumption estimation and optimization. In fact, we can estimate and optimize the power consumption of mobile devices from many angles including circuit, architecture, and system level studies. So far, there exist only a few studies pertaining to this aspect of green mobile communications. For example, in [114], the authors have presented tools and methods for collecting and analyzing logs of real activity patterns in order to characterize the power consumption and guide the optimization of mobile architectures. They demonstrate a saving of up to 10 percent system power with minimal impact on user satisfaction.

However, due to the support of multimedia services as well as the ubiquity of heterogeneous network activities, it has been quite challenging to predict and optimize energy consumption of the involved mobile devices. For example, considering a scenario in which an MS is able to connect to the Internet through either a WiFi AP, a cellular BS directly, or via an ad hoc relay, what would be the optimum energy consumption level for the MS battery in such a scenario?

In the following, we summarize the existing energy management techniques proposed for mobile devices and present the state-of-the-art techniques for MS energy consumption estimation, prediction, and optimization. Then, the two approaches proposed in Paper D for MS energy consumption prediction and optimization are outlined respectively.

5.2 A Retrospect on Energy Management Techniques for Mobile Devices

The area of energy conservation for mobile devices can be viewed from different perspectives considering protocols, battery energy measurements, models, etc.

5.2.1 Protocol-level View

There have been many endeavours for solving the problem of battery drain from protocol design point of view. In [115], the authors show that cellular networks present high tail energy overhead caused by the Forward Access Channel (FACH) in cellular networks. Tail energy is defined as the energy wasted when the system remains in high-power state even after the completion of a transfer. The authors present a protocol referred to as *TailEnder*, which is designed to save energy in mobile handsets by scheduling data transmissions using pre-fetching and caching for applications in order to minimize the tail energy caused by the inactivity timer [116].

Two more recent efforts for energy efficient management of power states of wireless networks are [117] and [118]. [117] utilizes short time intervals during wireless transmission in which it is possible to set the NIC to idle state. The authors demonstrate a reduction of about 30 percent power consumption for wireless transceivers with various applications without degrading user experience. The scheme in [118] reduces energy consumption by allowing the wireless interface to sleep during data transfers. This is achieved by shaping the traffic to combine small gaps between packets into longer sleep intervals during data transmissions. The authors report that the wireless interface can be allowed to sleep for about 40 percent of the time for a 10 MB data transfer.

Furthermore, the effects on energy efficiency of wireless networks from traffic patterns are also reported in the literature. For example, in [119] the authors describe how to reshape TCP traffic into periodic bursts with the same average throughput as the server transmission rate so that the client is able to accurately predict the packet arrival time and set its wireless interface to low-power mode accordingly.

Data compression techniques can also be used for reducing energy consumption of mobile devices. In [120], the authors show the possibility of more than 50 percent energy saving if the text data is compressed before transmission over a WiFi link. However, it is worth mentioning that such a technique requires extra energy consumption for data compression in runtime.

Another interesting study [65] takes the advantage of multiple supported wireless interfaces in a mobile phone to explore energy consumption reduction. The authors propose an energy management architecture in which the cellular radio wakes up the WLAN interface upon an incoming VoIP call to redirect the call through the WiFi interface, through which 70 $\sim$ 80 percent of energy saving is reported. A similar concept has, however, been proposed in [64] before [65].

It has also been observed that WiFi energy optimization has conventionally been designed with a single AP in mind. However, network contention among different APs can increase a client's energy consumption since each client may have to keep awake for longer duration before its own AP gets a chance to send its packets to it. In a very recent study [121], the authors design a system, known as *SleepWell*, which achieves energy efficiency by evading network contention. In their paper, the authors propose that the APs regulate the sleeping window of their clients in a manner that different APs are active/inactive during non-overlapping time windows.

5.2.2 Measurements and Models for Energy Consumption

The authors in [122] give a detailed analysis of energy consumption in mobile handsets by presenting a detailed breakdown of the power consumption of Openmoko Neo Freerunner phone (2.5G). The reason to use this device for their measurements is the free availability of its circuit details in contrast with other platforms such as Windows Mobile and iPhone. They conclude that the most energy consuming components in a mobile phone are the GSM module and the display panel, i.e., touchscreen, backlight, etc. For example, the GSM module consumes about 700 $\sim$ 800 mW during a phone call. They also show that the content in the screen may have huge effect on the power consumption of the screen, e.g., about 33 mW with white screen and about 74 mW with black screen. The results are validated and compared with Google Nexus and HTC Dream mobile phones. The authors of that paper also model five usage profiles and simulate the energy consumption per day for these profiles.

Furthermore, energy measurement of WiFi interfaces in smartphones is carried out in [123]. The authors present a platform to run energy measurements in mobile phones by using high-resolution power meters. The paper also analyzes the cost of sending messages over an IEEE 802.11 link and concludes that energy cost per transmitted kilobits varies with buffer size. More attempts to create general system-level power models can be found in [124] and [114].

In parallel to the experimental work presented so far in this section which relies mainly on measurement-based studies, theoretical studies are also ongoing. Some of the work in the direction of energy consumption prediction and optimization is summarized below.

5.2.3 MS Energy Consumption Prediction and Optimization

A modest amount of work exists on battery life prediction for MSs operated in a heterogeneous [125] [126] environment. For example, in [127], the authors present an energy optimization mechanism for MSs using vertical handoff between WLAN and CDMA2000 networks. The paper develops an algorithm to allow the MS to always get connected to the most cost-effective network. In [115], the authors build a protocol that reduces energy consumption of common mobile applications based on their energy consumption model. However, how their model predicts the long-term energy consumption in a heterogeneous environment is not discussed.

Moreover, [128] discusses the optimization of power consumption of applications which have significant effects on battery lifetime. The authors present synchronization algorithms for transmission of application updates and observe improvement in battery lifetime of the mobile terminals. Their focus is mainly on passive applications, which run as always-on background process in the mobile phones [129]. Another work [130] predicts the power consumption of always-on applications for mobile terminals operated in a CDMA network.

From the above literature survey, we observe that there is very little work on the energy consumption prediction and optimization for MSs operated in a heterogeneous scenario. This observation triggers our motivation for the work in Paper D.

5.3 Energy Consumption Prediction and Optimization in Heterogeneous Networks: Our Approach

In HNs, it is obvious and well-known that different network interfaces embedded in MSs consume power at different rates [131] (and references therein). Indeed the power consumption rate is dependent on the network to which the MS is connected to (e.g., depending on supported data rate, distance between transceivers etc). This is also dependent on the state of the MS, i.e., whether it is transmitting, receiving, or in idle state. Therefore, the power consumption of such MSs can be predicted and optimized. This can be achieved by, for example, configuring the MS to automatically select (or de-select) different available connections based on the power consumption, cost-effectiveness and user's preference. Given heterogeneous connections, a natural question that arises here is – what will be the battery lifetime of MSs operated in such environment and can we predict the asymptotic battery lifetime of such MSs? Paper D makes an initial effort to answer this question.

In our earlier work, i.e., Paper B, we present a distance-based power consumption analysis for MSs in a heterogeneous wireless network. Based on the same scenario viz an HN comprising of three different networks, i.e., a WLAN, an ad hoc, and a cellular network, we study energy consumption prediction and optimization from a different perspective in Paper D. More specifically, an energy consumption analysis involving Markov process is performed for an MS operated in an HN. The main task of Paper D is to predict and optimize the long-term battery energy consumption of such an MS.

5.3.1 Discrete-Time Markov Chain-based Energy Prediction

There are two methods proposed in Paper D. In the first part, a Discrete-Time Markov Chain (DTMC) with rewards is employed to analyze the asymptotic value of MS energy consumption given that it starts from a certain network link. For the DTMC analysis of the MS energy consumption, the HN is considered as a process where a state of the process refers to the link in which the MS can operate. Correspondingly, there are three states, i.e., W, A, and C, (denoted for the sake of expression convenience as states 1, 2, and 3) respectively, representing the MS operated in WLAN, Ad hoc, and Cellular link. A probability transition matrix, **P**, represents all the transition probabilities between states. Moreover, there is a corresponding cost/reward when the system makes (or does not make) a transition from one state to another. The cost matrix is represented by **R**. The reward (or the cost) in our study is considered as the MS energy consumption per communicated byte in the corresponding link. Therefore, for the DTMC, the expected total reward, $v_i(n)$, in n transitions if the process starts in state i, can be written as

$$v_i(n) = \sum_{j=1}^{N} p_{ij}[r_{ij} + v_j(n-1)] = q_i + \sum_{j=1}^{N} p_{ij} v_j(n-1), \tag{1}$$

where $i = 1,2,\ldots,N$ and $n = 1,2,\ldots$. p_{ij} and r_{ij} respectively represent the ijth element of matrices **P** and **R**. The product, q_i, of p_{ij} and r_{ij} is the expected reward for the next transition if the current state is i.

Based on the expected total reward in the next n transitions, the energy consumption of MS can be predicted because the difference between energy consumption for any two states tends to converge to a constant after a few initial transitions.

A relatively detailed description of this effect is explained in Paper D.

5.3.2 Finite Markov Decision Process for Energy Optimization

The second part of Paper D utilizes an FMDP with rewards. The state space in our FMDP contains all the possible states of the process while the action space is composed of all the possible actions that can be taken from these states. With certain probabilities, an action may (or may not) cause a state transition.

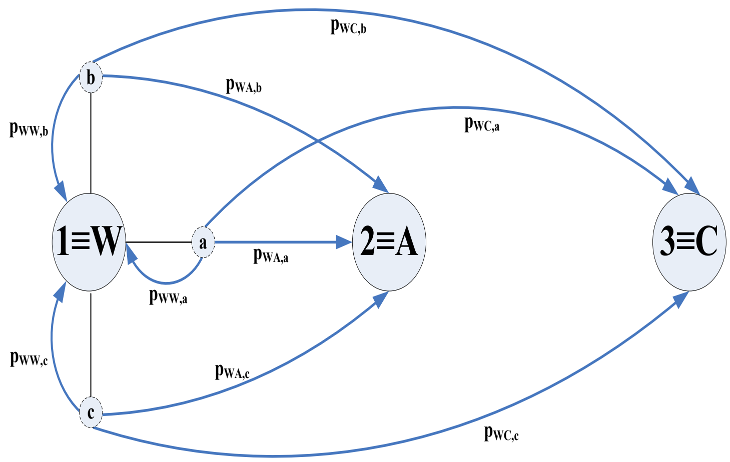


Figure 5.1: Partial FMDP for the MS with initial state at W.

A partial pictorial illustration of the FMDP utilized in Paper D is presented in Fig. 5.1. Here, the small dashed circles denote the action and the large solid circles denote the state of the system. For example, $p_{WA,c}$ means the probability of the system moving from state W to state A when action c is chosen. In the figure, to avoid pictorial complexity, all the possible transitions from only one state to other states are shown. The actual pictorial complexity of the process is in fact three-folds of the one shown in Fig. 5.1.

Three actions are defined in our FMDP, i.e., a, b, and c, respectively representing the MS operating in the High Energy Saving (HES), Low Energy Saving (LES), and Energy Ignorant (EI) modes. As shown in Fig. 5.1, an action may or may not lead the system to another state with certain probabilities.

A policy is a set of all possible actions that can be taken at time n in state i. In Paper D, in order to optimize the MS power consumption, we start with an artificial initial policy and utilize policy improvement methodology to arrive at the set of action-state pairs for each state that give the minimum energy consumption for the MS in the long-term of the system. For this purpose, Bellman's Principle of Optimality [132] [133] is adopted.

The numerical results suggest that by using our proposed approach, the MS can gradually converge to an optimized policy to ensure minimized overall long-term battery energy consumption. The convergence time is, however, dependent on the size of the action and state spaces. The results also reveal on how battery energy is consumed in heterogeneous networks.

5.4 Distinguishing Aspects of Paper D

There are two major distinguishing aspects in Paper D. First, a Markovian analysis with reward is proposed and presented for the energy consumption analysis of MS operated in an HN. This method enables us to predict the asymptotic battery lifetime of an MS in such networks.

A drawback of the DTMC-based approach is that the system has fixed transition directions. Therefore, the feasibility of multiple transition directions in order to minimize energy consumption is also explored. For this purpose, an FMDP is employed to provide more flexibility for the MS to select the least energy consuming alternative to ensure the overall long-term energy consumption minimization. The MS can progressively reach to an optimized policy to minimize its energy consumption. Furthermore, the convergence of the utilized policy iteration algorithm, which depends on the number of states and actions in the system, is also investigated. To the best of our knowledge, the FMDP-based approach for MS energy optimization is originally proposed in Paper D.

5.5 Chapter Summary

In this chapter, we gave a generalized retrospective overview on the efforts made in the direction of battery energy management as well as energy consumption prediction and optimization for mobile devices. Then we outlined respectively the two approaches originally proposed in Paper D for MS energy consumption prediction and optimization in HNs, i.e., the DTMC- and FMDP- based approaches.

Chapter 6

Concluding Remarks

In this final chapter, the scientific contributions of this dissertation are presented as a whole. The limitations of this thesis are also discussed. Finally, we point out a few future research directions in the area of energy conservation for MSs and wireless netwroks.

6.1 Summary of the Research Work and Major Scientific Contributions

The economic growth and environmental concerns have raised an imperative demand for green communications worldwide. The tasks of achieving green communications, however, need to be addressed from diverse perspectives. For instance, energy consumption reduction from network deployment to BS design, from operating systems to optimization algorithms, from MS architecture design to protocol and software implementation etc. In this dissertation, we attempt to address a few interesting energy conservation issues and propose our solutions.

Most of the techniques proposed for MS energy conservation are aimed at achieving improvements at different individual communication layers or designing energy efficient communication protocols. Multi-hop transmission has also been proposed as a promising solution, among many others. With this regard, the employment of relay nodes is able to reduce the energy consumption of the sending nodes, nevertheless, at an expense of the relay node's battery. Moreover, the problem of battery energy conservation becomes more complicated when the MS (or the relay node) is operated in a heterogeneous environment.

From the network perspective, many solutions have been proposed in the past decade for network energy conservation, achieving however only modest network energy conservation. Therefore, there is a need to explore more, both theoretical and practical, solutions which can save a reasonable amount of network energy.

In this dissertation, to reflect the goal of MS energy conservation in heterogeneous networks, methods using policy- and distance-based calculations are developed and investigated analytically. Furthermore, to reflect the goal of network-level energy conservation, solutions by switching off the most energy-consuming network components are proposed and analyzed mathematically. The objectives of the thesis have been achieved through comprehensive literature survey and in-depth studies in the identified field. Correspondingly, these achievements have resulted in a series of peer-reviewed international publication: two in journals and seven in conference proceedings. In the following, the main contributions of this dissertation are summarized.

- A continuous-space analytical approach for analyzing a hybrid network with uniform node distribution has been proposed and investigated in order to find network capacity and the optimal placement of the relay node.

- A distance-based MS energy computation and conservation approach is proposed in a heterogeneous wireless network. The possibilities of using distinct network links for uploading and downloading traffic are explored and compared. The numerical results suggest that different links can be used for uplink and downlink traffic in order to save MS battery energy. This work also explores the potential of energy-aware mobile-centric handover decision making in heterogeneous networks.

- An approach to switch off the middle cell BS in a three-cell network with pico- or micro-cells is proposed, based on the traffic density in the involved sectors. Two schemes, one deterministic and another one FMDP-based probabilistic, are presented and analyzed for this purpose. A policy to operate the network in the power-saving mode or normal operation is proposed, and the ping-pong effect is avoided based on the proposed hysteresis region. The two proposed schemes are studied in-depth and compared with each other in terms of network transmission power consumption.

- A DTMC-based and an FMDP-based methods for energy consumption prediction and optimization of an MS are proposed. The methodologies used in this work are applicable for hybrid mobile and wireless networks with heterogeneous links.

6.2 Limitations of the Research

Honestly speaking, there are certain limitations in the research scenarios as well as methodologies used in this dissertation. For example, in the problem of relay node placement, although relay node is utilized for coverage extension and thus battery conservation is achieved for the corresponding MS, it comes at the expense of extra battery energy consumption of the relay node itself. How to reduce this consumption while still achieving network extension with optimum capacity is an open question. For an MS operated in a heterogeneous network, although the utilization of different links for uplink and downlink traffic has been explored and shown to be of promising potential for MS energy conservation, how to select the optimum link still needs more exploration from protocol design point of view. Again, to select which link within the heterogeneous network in order to reduce the overall MS and relay node energy consumption needs further investigation by jointly considering multiple parameters such as link distance, path-loss coefficient, and data rate. Moreover, although the utilization of Markov-based analysis is useful in prediction and optimization of MS energy consumption, how to define a suitable policy in a multi-mobile scenario is an open question. Furthermore, joint optimization of data-transfer speed and MS energy consumption for heterogeneous traffic also needs to be investigated.

In the scenario when the middle cell is switched off, a major concern can be the interference at the sector edges since the same channel is used for the whole cell(s). This issue has not been addressed in this dissertation and needs therefore further attention.

6.3 Suggestions for Future Research

This dissertation is primarily aimed at exploring and proposing a few energy reduction methodologies for MS in heterogeneous networks. Another basic perspective of the thesis work is to investigate network-level energy conservation methodologies. Although the thesis work covers various aspects of the above mentioned issues, there are still numerous problems that remain to be tackled in the future. In the following, some of these directions are highlighted.

- Investigating the methods for efficient selection of a working relay node when multiple relay nodes are present within the coverage is a challenging but fertile research area. Another possible topic towards this direction is to explore methods to improve the overall optimum capacity for such a network.

- Handover strategies for an MS operating in a heterogeneous network exist in the literature. However, there is a need to develop optimum energy-aware mobile-centric handover schemes for such networks, by taking MS battery and other available network resources into account. The compatibility and comparison of such energy-aware handover schemes with the IEEE 802.21 standard is another interesting topic.

- The scenario considered in this thesis where a BS may be switched off for network energy saving is somewhat simplistic, consisting only three micro-cells. A consideration of multi-tiers of cells in order to achieve network energy reduction will for sure produce valuable research results, especially when inter- and intra-cell interference is considered. Extensions to analysis with asymmetric traffic distribution and non-uniform MS distribution also need to be explored in the future.

- Simulation-based experiments and real-life implementation of Markov chain-based probabilistic schemes for MS battery energy conservation and lifetime prediction in heterogeneous networks are appreciated.

- As the energy consumption of mobile nodes may vary according to distances, considering the mobility of MSs and relay nodes deserves further investigation.

REFERENCES

[1] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, "A Survey of Self Organization in Future Cellular Networks," *IEEE Communications Surveys and Tutorials*, vol. PP, no. 99, pp. 1–26, Feb. 2012.

[2] M. S. Kuran and T. Tugcu, "A Survey of Emerging Broadband Wireless Access Technologies," *Computer Networks*, vol. 51, no. 11, pp. 3013–3046, Aug. 2007.

[3] I. Chlamtac, M. Conti, and J. J.-N. Liu, "Mobile Ad Hoc Networking: Imperatives and Challenges," *Ad Hoc Networks*, vol. 1, no. 1, pp. 13–64, Jul. 2003.

[4] EU ENSURE Project, "Energy Savings in Urban Quarters through Rehabilitation and New Ways of Energy Supply (ENSURE)." Available at http://www.ensure-project.eu.

[5] Global Action Plan, "An Inefficient Truth," 2007. Available at http://www.globalactionplan.org.uk/.

[6] Alcatel Lucent Bell Labs, "Green Touch Consortium." Available at http://www.greentouch.org/.

[7] EU TREND Project, "Towards Real Energy-efficient Network Design (TREND)." Available at http://www.fp7-trend.eu/.

[8] Nokis Siemens Networks, "Go Green, Save Energy, Win Customer Kudos," *Unite*, no. 7, pp. 16–17, Feb. 2010.

[9] Mobile Virtual Centre of Excellence, "Core 5 Research Program." Available at http://www.mobilevce.com/frames.htm?core5research.htm.

[10] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell Zooming for Cost-efficient Green Cellular Networks," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 74–79, Nov. 2010.

[11] H. Zhang, A. Gladisch, M. Pickavet, Z. Tao, and W. Mohr (guest editors), "Special Issue on Energy Efficiency in Communications: Part II," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 30–82, Jun. 2011.

[12] V. Mancuso and S. Alouf, "Reducing Costs and Pollution in Cellular Networks," *IEEE Communications Magazine*, vol. 49, no. 8, pp. 63–71, Aug. 2011.

[13] I. Ashraf, F. Boccardi, and L. Ho, "Sleep Mode Techniques for Small Cell Deployments," *IEEE Communications Magazine*, vol. 49, no. 8, pp. 72–79, Aug. 2011.

[14] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. S. Thompson, I. Ku, C.-X. Wang, T. A. Le, M. R. Nakhai, J. Zhang, and L. Hanzo, "Green Radio: Radio Techniques to Enable Energy-efficient Wireless Networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 46–54, Jun. 2011.

[15] M. A. Marsan and M. Meo, "Energy Efficient Wireless Internet Access with Cooperative Cellular Networks," *Computer Networks*, vol. 55, no. 2, pp. 386–398, Feb. 2011.

[16] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal Energy Savings in Cellular Access Networks," in *Proc. First International Workshop on Green Communications*, Dresden, Germany, Jun. 2009.

[17] K. Dufkov, M. Bjelica, B. Moon, L. Kencl, and J.-Y. L. Boudec, "Energy Savings for Cellular Networks with Evaluation of Impact on Data Traffic Performance," in *Proc. European Wireless*, Lucca, Italy, Apr. 2010.

[18] A. J. Fehske, F. Richter, and G. P. Fettweis, "Energy Efficiency Improvements through Micro Sites in Cellular Mobile Radio Networks," in *Proc. IEEE 2nd International Workshop on Green Communications, in conjunction with GLOBECOM 2009*, Honolulu, HI, USA, Nov.-Dec. 2009.

[19] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Switch-off Transients in Cellular Access Networks with Sleep Modes," in *4th Workshop on Green Communications, in conjunction with IEEE ICC*, Kyoto, Japan, Jun. 2011.

[20] A. P. Jardosh, K. Papagiannaki, E. Belding, K. C. Almeroth, G. Iannaccone, and B. Vinnakota, "Green WLANs: On-demand WLAN Infrastructures," *Mobile Networks and Applications, SI on Recent Advances in WLANs*, vol. 14, no. 6, pp. 798–814, Apr. 2009.

[21] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "A Simple Analytical Model for the Energy-efficient Activation of Access Points," in *Proc. e-Energy 2010*, Passau, Germany, Apr. 2010.

[22] V. Baiamonte and C.-F. Chiasserini, "Saving Energy During Channel Contention in 802.11 WLANs," *Mobile Network Applications*, vol. 11, no. 2, pp. 287–296, Apr. 2006.

[23] S. Sayed, Y. Yang, H. Guo, and H. Hu, "Analysis of Energy Efficiency of a Busy Tone Based Cooperative MAC Protocol for Multi-rate WLANs," in *Proc. IEEE WCNC*, Sydney, Australia, Apr. 2010.

[24] X. Perez-Costa and D. Camps-Mur, "IEEE 802.11e QoS and Power Saving Features Overview and Analysis of Combined Performance," *IEEE Wireless Communications*, vol. 17, no. 4, pp. 88–96, Aug. 2010.

[25] E. H. Ong, K. Mahata, and J. Y. Khan, "Energy Efficient Architecture for Green Handsets in Next Generation IP-based Wireless Networks," in *Proc. IEEE ICC*, Kyoto, Japan, Jun. 2011.

[26] L. Feeney, B. Ahlgren, and A. Westerlund, "Spontaneous Networking: An Application-oriented Approach to Ad Hoc Networking," *IEEE Communications Magazine*, vol. 39, no. 6, pp. 176–181, Jun. 2001.

[27] M. S. Corson, J. P. Maker, and J. H. Cernicione, "Internet-based Mobile Ad Hoc Networking," *IEEE Internet Computing*, vol. 3, no. 4, pp. 63–70, Jul.–Aug. 1999.

[28] S. Giordano, *Mobile Ad-hoc Networks in Handbook of Wireless Networks and Mobile Computing*. I. Stojmenovic ed., Wiley, New York, USA, 2002.

[29] S. Sarkar and K. N. Sivarajan, "Fairness in Cellular Mobile Networks," *IEEE Trans. Inform. Theory*, vol. 48, no. 8, pp. 2418–2426, Aug. 2002.

[30] H. Wu, C. Qiao, S. De, and O. Tonguz, "Integrated Cellular and Ad Hoc Relaying System: iCAR," *IEEE J. Select. Areas Commun.*, vol. 19, no. 10, pp. 2105–2115, Oct. 2001.

[31] H. Luo, R. Ramjee, P. Sinha, L. E. Li, and S. Lu, "UCAN: A Unified Cellular and Ad Hoc Network Architecture," in *Proc. ACM MOBICOM*, San Diego, CA, USA, Sept. 2003.

[32] H. Y. Hsieh and R. Sivakumar, "On Using the Ad Hoc Network Model in Cellular Packet Data Networks," in *Proc. ACM MOBIHOC*, Lausanne, Switzerland, Jun. 2002.

[33] 3GPP Project, "Third Generation Partnetship Project." Available at http://www.3gpp.org/lte-advanced.

[34] D. Soldani and S. Dixit, "Wireless Relays for Broadband Access," *IEEE Communications Magazine*, vol. 46, no. 3, pp. 58–66, Mar. 2008.

[35] B. Liu, Z. Liu, and D. Towsley, "On the Capacity of Hybrid Wireless Networks," in *Proc. IEEE INFOCOM*, San Francisco, CA, USA, Mar.-Apr. 2003.

[36] Y. D. Lin and Y. C. Hsu, "Multihop Cellular: A New Architecture for Wireless Communications," in *Proc. IEEE INFOCOM*, Tel Aviv, Israel, Mar. 2000.

[37] S. Lee, S. Benerjee, and B. Bhattacharjee, "The Case for a Multihop Wireless Local Area Network," in *Proc. IEEE INFOCOM*, Hong Kong, China, Mar. 2004.

[38] R. Chakravorty, S. Agarwal, S. Banerjee, and I. Pratt, "MoB: A Mobile Bazaar for Wire-area Wireless Services," in *Proc. ACM MOBICOM*, Cologne, Germany, Aug. 2005.

[39] H. Viswanathan and S. Mukherjee, "Performance of Cellular Networks with Relays and Centralized Scheduling," in *Proc. IEEE VTC Fall 2003*, Orlando, FL, USA, Oct. 2003.

[40] A. Agarwal and R. R. Kumar, "Capacity Bounds for Ad Hoc and Hybrid Wireless Networks," *ACM Computer Comm. Rev.*, vol. 34, no. 3, pp. 71–81, Jul. 2004.

[41] S. Toumpis, "Capacity Bounds for Three Classes of Wireless Networks," in *Proc. ACM MOBIHOC*, Tokyo, Japan, May 2004.

[42] J. Cho and Z. J. Haas, "On the Throughput Enhancement of the Downstream Channel in Cellular Radio Networks through Multihop Relaying," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 1206–1219, Sept. 2004.

[43] S. Mukherjee and H. Viswanathan, "Analysis of Throughput Gains from Relays in Cellular Networks," in *Proc. IEEE GLOBECOM*, San Francisco, CA, USA, Dec. 2003.

[44] H. Zhai and Y. Fang, "Physical Carrier Sensing and Spatial Reuse in Multirate and Multihop Wireless Ad Hoc Networks," in *Proc. IEEE INFOCOM*, Barcelona, Spain, Apr. 2006.

[45] S. Misra, S. D. Hong, G. Xue, and J. Tang, "Constrained Relay Node Placement in Wireless Sensor Networks to Meet Connectivity and Survivability Requirements," in *Proc. IEEE INFOCOM*, Phoenix, AZ, USA, Apr. 2008.

[46] L. K. Law, S. V. Krishnamurthy, and F. Faloutsos, "Capacity of Hybrid Cellular-Ad Hoc Data Networks," in *Proc. IEEE INFOCOM*, Phoenix, AZ, USA, Apr. 2008.

[47] M. F. M. Hossain, A. Mammela, and H. Chowdhury, "Impact of Mobile Relays on Throughput and Delays in Multihop Cellular Networks," in *Proc. ICWMC*, Athens, Greece, Jul. 2008.

[48] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 388–404, Mar. 2000.

[49] TNS, "TNS Global Market Research: TNS User Study." Available at http://www.tnsglobal.com.

[50] K. Pentikousis, "In Search of Energy-efficient Mobile Networking," *IEEE Communications Magazine*, vol. 48, no. 1, pp. 95–103, Jan. 2010.

[51] M. Armand and J.-M. Tarascon, "Building Better Batteries," *Nature*, vol. 451(7179), pp. 652–657, Feb. 2008.

[52] J. Baker, "New Technology and Possible Advances in Energy Storage," *Energy Policy*, vol. 36, no. 12, pp. 4368–4373, 2008.

[53] P. J. Hall and E. Bain, "Energy-storage Tecnhologies and Electricity Generation," *Energy Policy*, vol. 36, no. 12, pp. 4352–4355, 2008.

[54] IEEE 802.11 WG, "IEEE Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications." IEEE Standard, 2007.

[55] Agere Systems, “WaveLAN 802.11b Chipset for Standard Form Factors.” Available at http://www.lsi.com/.

[56] S.-T. Cheng and M. Wu, “Contention-polling Duality Coordination Function for IEEE 802.11 WLAN Family,” *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 779–788, Mar. 2009.

[57] M. S. Gast, *802.11 Wireless Networks The Definitive Guide*. O’Reilly and Associates Inc., CA, USA, 2nd ed., 2005.

[58] V. Jacobson, “A New Way to Look at Networking,” Aug. 2006. Available at http://video.google.com/videoplay?docid=-6972678839686672840.

[59] Y. Neuvo, “Cellular Phones as Embedded Systems,” in *Proc. IEEE ISSCC*, San Fransisco, CA, USA, Feb. 2004.

[60] R. Wang, J. Tsai, C. Maciocco, T.-Y. C. Tai, and J. Wu, “Reducing Power Consumption for Mobile Platform via Adaptive Traffic Coalescing,” *IEEE J. Select. Areas Commun.*, vol. 29, no. 8, pp. 1618–1629, Sept. 2011.

[61] D. Qiao and K. Shin, “Smart Power-saving Mode for IEEE 802.11 Wireless LANs,” in *Proc. IEEE INFOCOM*, Miami, FL, USA, Mar. 2005.

[62] V. Namboodiri and L. Gao, “Towards Energy Efficient VoIP over Wireless LANs,” in *Proc. ACM MOBIHOC*, Hong Kong, China, May 2008.

[63] S. Makineni, R. Iyer, P. Sarangam, D. Newell, L. Zhao, R. Illikkal, and J. Moses, “Receive Side Coalescing for Accelerating TCP/IP Processing,” in *Proc. High Performance Computing, HiPC’06*, Bangalore, India, Dec. 2006.

[64] E. Shih, P. Bahl, and M. Sinclair, “Wake on Wireless: An Event Driven Energy Saving Strategy for Battery Operated Devices,” in *Proc. ACM MOBICOM*, Atlanta, GA, USA, Sept. 2002.

[65] T. Agarwal, S. Hodges, J., R. Chandra, P. Bahl, and R. Gupta, “Somniloquy: Augmenting Network Interfaces to Reduce PC Energy Usage,” in *Proc. the USENIX Symposium on Networked Systems Design and Implementation (NSDI 09)*, Boston, MA, USA, Apr. 2009.

[66] A. Kulkarni, R. Wang, C. Macioicco, S. Bakshi, and J. Tsai, “An Energy Efficient Communication Scheme for Connected Mobile Platforms,” in *Proc. IEEE ICC*, Dresden, Germany, Jun. 2009.

[67] A. Chandrakasan and R. W. Brodersen, "Minimizing Power Consumption in Digital CMOS Circuits," *Proceedings of the IEEE*, vol. 83, no. 4, pp. 498–523, Apr. 1995.

[68] R. Gonzalez and M. Horowitz, "Energy Dissipation in General Purpose Microprocessors," *IEEE J. Solid-State Circuits*, vol. 31, no. 9, pp. 1277–1284, Sept. 1996.

[69] IEEE 802.21 WG, "IEEE Standard for Local and Metropolitan Area Networks - Media Independent Handover Services." IEEE Standard, 2008.

[70] A. de la Oliva, T. Malia, A. Vidal, C. Bernardos, I. Soto, and A. Banchs, "A Case Study: IEEE 802.21 Enabled Mobile Terminal for Optimized WLAN/3G Handovers," *ACM SIGMOBILE Mobile Computing and Communication Review*, vol. 11, no. 2, pp. 29–40, Apr. 2007.

[71] A. Dutta, S. Das, D. Famolari, Y. Ohba, K. Taniuchi, T. Kodama, and H. Schulzrinne, "Seamless Handover Across Heterogeneous Networks - An IEEE 802.21 Centric Approach," in *Proc. WPMC*, Aalborg, Denmark, Sept. 2005.

[72] P. Henrik, "Energy-aware Network Selection using Traffic Estimation," in *Proc. 1st ACM Workshop on Mobile Internet through Cellular Networks*, Beijing, China, Sept. 2009.

[73] T. Pering, Y. Agarwal, R. Gupta, and R. Want, "Reducing the Power Consumption of Wireless Mobile Devices with Multiple Radio Interfaces," in *Proc. MOBISYS*, New York, USA, Jun. 2006.

[74] A. Rahmati and L. Zhong, "Context for Wireless: Context-sensitive Energy-efficient Wireless Data Transfer," in *Proc. MOBISYS*, San Juan, Puerto Rico, Jun. 2007.

[75] C. Levy-Bencheton and G. Villemaud, "Power Consumption Optimization in Multi-mode Mobile Relays," in *Proc. 12th European Conference on Wireless Technology*, Rome, Italy, Sept. 2009.

[76] E. Shih, S. Cho, N. Ickes, R. Min, A. Sinha, A. Wang, and A. P. Chandrakasan, "Physical Layer Driven Algorithm and Protocol Design for Energy-efficient Wireless Sensor Networks," in *Proc. ACM MOBICOM*, Rome, Italy, Jul. 2001.

[77] F. Richter and G. P. Fettweis, "Cellular Mobile Network Densification Utilizing Micro Base Stations," in *Proc. IEEE ICC*, Cape Town, South Africa, May 2010.

[78] X. Wu, B. Murherjee, and D. Goshal, "Hierarchical Architectures in the Third-generation Cellular Networks," *IEEE Wireless Communication Magazine*, vol. 11, no. 3, pp. 62–71, Jun. 2004.

[79] H. Claussen, L. T. W. Ho, and L. G. Samuel, "Financial Analysis of a Picocellular Home Network Deployment," in *Proc. IEEE ICC*, Glasgow, Scotland, Jun. 2007.

[80] S. Bhaumik, G. Narlikar, and S. Chattopadhay, "Breath to Stay Cool: Adjusting Cell Sizes to Reduce Energy Consumption," in *Proc. 1st ACT SIGCOMM Workshop on Green Networking*, New Delhi, India, Aug.-Sept. 2010.

[81] A. Amanna, "Green Communications: Annotated Review and Research Vision." Technical Report, Virginia Tech. 2010.

[82] H. Claussen, L. T. W. Ho, and F. Pivit, "Effects of Joint Macrocell and Residential Picocell Deployment on the Network Energy Efficiency," in *Proc. IEEE PIMRC*, Cannes, France, Sept. 2008.

[83] J. Kim and B. Kim and Y. Y. Woo, "Advanced Design of Linear Doherty Amplifier for High Efficiency using Saturation Amplifier," in *Proc. IEEE/MTT-S International Microwave Symposium*, Honolulu, HI, USA, Jun. 2007.

[84] J. Wigard, T. Kolding, L. Dalsgaard, and C. Coletti, "On the User Performance of LTE UE Power Saving Schemes with Discontinuous Reception in LTE," in *Proc. Communications Workshop 2009, in conjunction with IEEE ICC 2009*, Dresden, Germany, Jun. 2009.

[85] L. Zhou, H. Xu, H. Tian, Y. Gao, L. Du, and L. Chen, "Performance Analysis of Power Saving Mechanism with Adjustable DRX Cycles in 3GPP LTE," in *Proc. IEEE VTC 2008-Fall*, Calgary, Canada, Sept. 2008.

[86] R. Gupta and E. C. Strinati, "Green Scheduling to Minimize Base Station Transmit Power and UE Circuit Power Consumption," in *Proc. IEEE PIMRC*, Toronto, Canada, Sept. 2011.

[87] D. Ezri, S. Shilo, B. Benmoshe, and E. Berliner, "Performance Study of Green Cellular - An Architecture for Minimal Emission from Mobile Stations," in *Proc. IEEE PIMRC*, Istanbul, Turkey, Sept. 2010.

[88] 4G Americas, “The Benefits of SON in LTE: Self-optimizing and Self-organizing Networks.” White Paper, Dec. 2009. Available at http://www.4gamericas.org/.

[89] L. C. Schmelz, J. L. Van Den Berg, R. Litjens, K. Zetterberg, M. Amirijoo, K. Spaey, I. Balan, N. Scully, and S. Stefanski, “Self-organization in Wireless Networks User Cases and Their Interrelation,” in *Proc. WWRF 22nd Meeting*, Paris, France, May 2009.

[90] T. Han and N. Ansari, “ICE: Intelligent Cell BrEathing to Optimize the Utilization of Green Energy,” *IEEE Communications Letters*, vol. 16, no. 6, pp. 866–869, Jun. 2012.

[91] L. Cai, H. Poor, Y. Liu, T. Luan, X. Shen, and J. Mark, “Dimensioning Network Deployment and Resource Management in Green Mesh Networks,” *IEEE Communications Magazine*, vol. 18, no. 5, pp. 58–65, Oct. 2011.

[92] I. Viering, M. Peltomaki, O. Trikkonen, M. Alava, and R. Waldhauser, “A Distributed Power Saving Algorithm for Cellular Networks,” in *Proc. IW-SOS*, Zurich, Switzerland, Dec. 2009.

[93] K. Samdanis, D. Kutscher, and M. Brunner, “Self-organizing Energy Efficient Cellular Networks,” in *Proc. IEEE PIMRC*, Istanbul, Turkey, Sept. 2010.

[94] K. Samdanis, D. Kutscher, and M. Brunner, “Dynamic Energy-aware Network Re-configuration for Cellular Urban Infrastructure,” in *Proc. IEEE GLOBECOM Workshops*, Anaheim, CA, USA, Dec. 2010.

[95] S. H. Shin and K. S. Kwak, “Power Control for CDMA Macro-micro Cellular System,” in *Proc. IEEE VTC Fall 2000*, Barcelona, Spain, Sept. 2000.

[96] H. Leem, Y. Baek, and D. K. Sung, “The Effects of Cell Size on Energy Saving, System Capacity, and Per-energy Capacity,” in *Proc. IEEE WCNC*, Shanghai, China, Apr. 2010.

[97] R. Pabst, B. H. Walke, D. C. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. D. Falconer, and G. P. Fettweis, “Relay-based Deployment Concepts for Wireless and Mobile Broadband Radio,” *IEEE Communications Magazine*, vol. 42, no. 9, pp. 80–89, Sept. 2004.

[98] X. J. Li, B. C. Seet, and P. H. J. Chong, "Multihop Cellular Networks: Technology and Economics," *Computer Networks*, vol. 52, no. 9, pp. 1825–1837, Jun. 2008.

[99] J. N. Laneman and G. W. Wornell, "Energy-efficient Antenna Sharing and Relaying for Wireless Networks," in *Proc. IEEE WCNC*, Chicago, IL, USA, Sept. 2000.

[100] A. Radwan and H. S. Hassanein, "NXG04-3: Does Multi-hop Communication Extend the Battery Life of Mobile Terminals," in *Proc. IEEE GLOBECOM*, San Francisco, CA, USA, Nov.-Dec. 2006.

[101] R. Rost and G. Fettweis, *Green Communication in Cellular Network with Fix Relay Nodes (Book Chapter).* Cambridge University Press, May 2011.

[102] A. Sendonaris, E. Erkip, and B. Aazhang, "User Cooperation Diversity, Part I: System Description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[103] M. Nokleby and B. Aazhang, "User Cooperation for Energy-efficient Cellular Communications," in *Proc. IEEE ICC*, Cape Town, South Africa, May 2010.

[104] F. Richter, A. J. Fehske, and G. P. Fettweis, "Energy Efficiency Aspects of Base Station Deployment Strategies for Cellular Networks," in *Proc. IEEE 70th VTC Fall*, Anchorage, AK, USA, Sept. 2009.

[105] D. Callin, H. Claussen, and H. Uzunalioglu, "On Femto Deployment Architectures and Macrocell Offloading Benefits in Joint Macro-femto Deployments," *IEEE Communications Magazine*, vol. 48, no. 1, pp. 26–32, Jan. 2010.

[106] J. Hoydis, M. Kobayashi, and M. Debbah, "Green Small-cell Networks," *IEEE Vehicular Technology Magazine*, vol. 6, no. 1, pp. 37–43, Mar. 2011.

[107] I. Ashraf, L. T. W. Ho, and H. Claussen, "Improving Energy Efficiency of Femtocell Base Stations via User Activity Design," in *Proc. IEEE WCNC*, Sydney, Australia, Apr. 2010.

[108] R. Wang, J. S. Thompson, H. Haas, and P. M. Grant, "Sleep Mode Design for Green Base Stations," *IET Communications*, vol. 5, no. 18, pp. 2606–2616, Dec. 2011.

[109] M. Gupta and S. Singh, "Greening of the Internet," in *Proc. ACM SIGCOMM*, Karlsruhe, Germany, Aug. 2003.

[110] L. Saker, S.-E. Elayoubi, and T. Chahed, "Minimizing Energy Consumption via Sleep Mode in Green Base Stations," in *Proc. IEEE WCNC*, Shanghai, China, Apr. 2010.

[111] M. A. Marsan and M. Meo, "Energy Efficient Management of Two Cellular Access Networks," in *Proc. GreenMetrics 2009 Workshop*, Seattle, WA, USA, Jun. 2009.

[112] Z. H. Abbas and F. Y. Li, "A Novel Teletraffic-based Power Saving Scheme for Cellular Networks with Microcells," in *Proc. 7th IEEE Performance and Management of Wireless and Mobile Networks (P2MNET) Workshop, in conjunction with 36th IEEE LCN*, Bonn, Germany, Oct. 2011.

[113] F. S. Hillier and G. J. Lieberman, *Introduction to Operations Research*. McGraw-Hill Companies Inc. New York, USA, 8th ed., 2005.

[114] A. Shye, B. Scholbrock, and G. Memik, "Into the Wild: Studying Real User Activity Patterns to Guide Power Optimizations for Mobile Architectures," in *Proc. IEEE/ACM International Symposium on Microarchitecture*, New York, USA, Dec. 2009.

[115] N. Balasubramanian, A. Balasubramanian, and A. Venkataramani, "Energy Consumption in Mobile Phones: A Measurement Study and Implications for Network Applications," in *Proc. 9th ACM SIGCOMM*, Chicago, IL, USA, Nov. 2009.

[116] C.-C. Lee, J.-H. Yeh, and J.-C. Chen, "Impact of Inactivity Timer on Energy Consumption in WCDMA and CDMA2000," in *Proc. Wireless Telecommunications Symposium*, Cal Poly Pomona Pomona, CA, USA, May 2004.

[117] J. Liu and L. Zhong, "Micro Power Management of Active 802.11 Interfaces," in *Proc. ACM MOBISYS*, Breckenridge, CO, USA, Jun. 2008.

[118] F. R. Dogar, P. Steenkiste, and K. Papagoannaki, "Catnap: Exploiting High Bandwidth Wireless Interfaces to Save Energy for Mobile Devices," in *Proc. ACM MOBISYS*, San Francisco, CA, USA, Jun. 2010.

[119] E. Tan, L. Guo, S. Chen, and X. Zhang, "PSM-throttling: Minimizing Energy Consumption for Bulk Data Communications in WLANs," in *Proc. IEEE International Conference on Network Protocols*, Beijing, China, Oct. 2007.

[120] K. C. Barr and K. Asanovic, "Energy-aware Lossless Data Compression," *ACM Transactions on Computer Systems*, vol. 24, no. 3, pp. 250–291, Aug. 2006.

[121] J. Manweiler and R. R. Choudhury, "Avoiding the Rush Hours: WiFi Energy Management via Traffic Isolation," *IEEE Trans. on Mobile Computing*, vol. 11, no. 5, pp. 739–752, May 2012.

[122] A. Carroll and G. Heiser, "An Analysis of Power Consumption in a Smartphone," in *Proc. USENIX Annual Technical Conference*, Boston, MA, USA, Jun. 2010.

[123] A. Rice and S. Hay, "Decomposing Power Measurements for Mobile Devices," in *Proc. IEEE PERCOM*, Mannheim, Germany, Mar.-Apr. 2010.

[124] Y. Xiao, R. Bhaumik, Z. Yang, M. Siekkinen, P. Savolainen, and A. Yla-Jaaski, "A System-level Model for Runtime Power Estimation on Mobile Devices," in *Proc. IEEE/ACM International Conference on Green Computing and Communications and International Conference on Cyber, Physical and Social Computing*, Hangzhou, China, Dec. 2010.

[125] M. Buddhikot, G. Chandranmenon, S. Han, Y. W. Lee, S. Miller, and L. Salgarelli, "Integration of 802.11 and third-Generation Wireless Data Networks," in *Proc. IEEE INFOCOM*, San Franciso, CA, USA, Mar.-Apr. 2003.

[126] M. Buddhikot, G. Chandranmenon, S. Han, Y. W. Lee, S. Miller, and L. Salgarelli, "Design and Implementation of a WLAN/CDMA2000 Interworking Architecture," *IEEE Communications Magazine*, vol. 41, no. 11, pp. 90–100, Nov. 2003.

[127] N. Salawu and E. N. Onwuka, "Energy Optimization Mechanism for Mobile Terminals using Vertical Handoff between WLAN and CDMA2000 Networks," *Leonardo Electronic Journal of Practices and Technologies*, pp. 51–58, Jul.–Dec. 2009.

[128] V. Kononen and P. Paakkonen, "Optimizing Power Consumption of Always-on Applications Based on Timer Alignment," in *Proc. 3rd International Conference on Communication Systems and Networks*, Bangalore, India, Jan. 2011.

[129] 3GPP Project, "Radio Resource Control (RCC) Protocol Specification." 3GPP, Tech. Rep. TS 25.331, Mar. 2009.

[130] H. Haverinen, J. Siren, and P. Eronen, "Energy Consumption of Always-on Applications in WCDMA Networks," in *Proc. VTC-Spring*, Dublin, Ireland, Apr. 2007.

[131] L. M. Feeney and M. Nilsson, "Investigating the Energy Consumption of a Wireless Network Interface in an Ad Hoc Networking Environment," in *Proc. IEEE INFOCOM*, Anchorage, AK, USA, Apr. 2001.

[132] R. Bellman, "The Theory of Dynamic Programming," *Bulliten of the American Mathematical Society*, pp. 503–515, Nov. 1954.

[133] M. Sniedovich, "A New Look at Bellman's Principle of Optimality," *Journal of Optimization Theory and Applications*, vol. 49, no. 1, pp. 161–176, Apr. 1986.

Part II

Paper A

A Continuous-space Analytical Approach for Relay Node Placement in Hybrid Cellular and Ad Hoc Networks

Title:	A Continuous-space Analytical Approach for Relay Node Placement in Hybrid Cellular and Ad Hoc Networks
Authors:	Ziaul Haq Abbas and Frank Y. Li
Affiliation:	Department of Information and Communication Technology, University of Agder (UiA), N-4898 Grimstad, Norway
Conference:	*IEEE International Symposium on Wireless Communication Systems, ISWCS 2009*, Siena, Italy, September 2009.
Copyright©:	IEEE

A Continuous-space Analytical Approach for Relay Node Placement in Hybrid Cellular and Ad Hoc Networks

Ziaul Haq Abbas and Frank Y. Li
Department of Information and Communication Technology
University of Agder (UiA), N-4898 Grimstad, Norway
E-mails: {ziaul.h.abbas, frank.li}@uia.no

Abstract — **A hybrid network is composed of a cellular component and an ad hoc component connected by a relay node, for the purpose of coverage extension and/or capacity improvement. In this paper, we analyze the capacity of such a hybrid network by employing a continuous-space analytical methodology based on circular geometry for uniformly distributed nodes. To achieve maximal overall capacity, the relay node needs to be placed in an optimum location between the base station and the mobile station located at the boundary of the hybrid network. Numerical results show that for obtaining the optimum overall capacity for the hybrid network, the placement of the relay node should be in a range which is neither too close nor too far away from the base station. For a given node density and path-loss coefficient, a precise location for relay node placement to achieve maximum overall capacity can be found using the presented method.**

I. Introduction

Tremendous amount of interest from the research community has been raised in recent years to enhance the coverage of the mobile cellular networks while maintaining better link quality and low cost. The improvement in the capacity of such networks has also been an active research topic. While there have been many attempts for increasing the cellular range (similarly for the capacity), there are practical and/or natural fundamental limits [1] on these quantities. For example, transmission power of a Base Station (BS) can not be increased beyond the limit defined by regulations. Similarly, the transmission power of a Mobile Station (MS) has to be kept within an optimum level due to, for example, interference and battery life considerations. When the coverage area of a BS becomes larger, interference increases due to higher number of interfering nodes, leading to reduced capacity and lower supported user data rate. By reducing the coverage area of the BS, the benefit of spatial reuse of spectrum is achieved. However, this benefit has to be traded off with the installation cost of a BS. As a consequence, the limits on the number of BSs in an area may create uncovered geographical spots. Moreover, the quality of signal in a circular region near a BS is better than in the far region (assuming omnidirectional antennas). Thus, the question of how to improve and extend the coverage (as well as capacity) of the network in the far region from the BS remains valid.

On the other hand, a good amount of research work has been carried out by the scientific community in the field of mobile ad hoc networks [2]. In an ad hoc network, two MSs can communicate with each other, either directly or via multihop relay, but the range of communication between two neighboring nodes is relatively short compared with the range between an MS and the BS in cellular networks. Some other advantages of ad hoc networks are that they can be easily deployed in an infrastructure-less manner and are of economically low cost.

A natural question is what happens if we integrate both the pure cellular and pure ad hoc network to create a Hybrid Network (HN) [3]. The low coverage but high capacity of ad hoc networks and the high coverage but low capacity of cellular networks can be traded off with each other to reach an optimum solution. This may lead to overall improvement in the cellular coverage as well as optimum capacity of the HN. Many such attempts have been made within this topic and relevant directions, for example [4].

In [5], the authors evaluated a hybrid network by reducing the coverage of a BS and relaying the traffic outside this coverage area. They intended to determine whether this scheme could improve the total performance in terms of capacity. Hexagonal geometry was used for this purpose and the analysis was carried out with r*egular* placement of the BS and users. Fixed spectral bandwidth was divided between the two components of the hybrid network using a weight coefficient. However, *how to find the optimum capacity of the HN has not been addressed in their work*. In [6], the authors studied the impact of mobile relays on throughput and delay. Hexagonal grid was used around the BS and the mobile stations were distributed randomly in a certain region. However, *in their scenario, the relay nodes are preselected closest to the BS*.

In this paper, we analyze the overall capacity of an HN in which a one-hop ad hoc network uses an MS from the cellular part of the HN as a relay node. The goal of this work is to find a position for an *optimal placement* of the relay node. By optimal placement, it is meant that an optimum location for the relay node is determined such that the maximum overall capacity is achieved in the HN. We also study how to find this overall capacity with respect to the change in node density and path-loss coefficient. A salient feature of our method which distinguishes this work from other existing approaches is that it is based on a continuous-space analytical approach with circular geometry while most other related work is based on discrete-space approach.

The rest of the paper is organized as follows. In Section II we describe the scenario of the studied network and present basic assumptions for the analysis which is given in Section III. Section IV presents and discusses the numerical results obtained through our method. Section V concludes the paper.

II. Hybrid Network Scenario and Assumptions

In this section, we describe the network architecture we are considering in an HN. Fig. A.1 shows the pictorial representation of the scenario studied in this paper. The overall region is divided into three sub-regions, namely, region-A, region-AB, and region-B. The values of R_1, R_2 and R_3 are not fixed so that the areas of all the three mentioned sub-regions also vary accordingly; and the distance from the BS to the MS located at the boundary of the region-B is $D = R_1 + R_2 + R_3$.

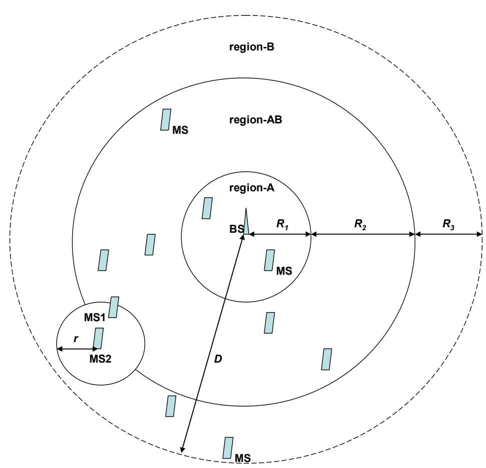


Figure A.1: Hybrid network architecture.

In our HN architecture, the MSs in region-A access the BS directly, and are referred to as *cellular nodes*. The mobile stations in region-B can also be reached by the BS, but with lower capacity. However, if a relay node within region-AB is used to forward packets then an MS in region-B may reach the BS with higher capacity. In other words, an MS in region-B, referred to as *ad hoc node*, can use

MSs in region-AB as relay nodes to reach the BS. Hence, MSs in region-AB are referred to as *relay nodes*, and may operate either in cellular or ad hoc mode.

We assume that each MS is equipped with dual radio interface, i.e., a cellular interface and an ad hoc interface. The operating frequency for an ad hoc node is different from the operating frequency of a cellular node. At any time instant, an MS can use one interface only. Each MS has a buffer of sufficient size to store data coming on its one interface and it may then relay the data through the same or the other interface (depending on which mode it is working on). A relay node may send its own data as well. We further assume that there is negligible processing delay at a node and no incoming packets are dropped.

In the considered network, all MSs are distributed uniformly. By using a kind of pilot signal, the BS can inform the MSs in the whole network about their working mode at any given instant either directly or through the relay nodes. The algorithm for updating the network topology information is assumed to be working on the backend.

For the HN, identical transmission power is assumed for all MSs so that the Signal-to-Interference Ratio (SIR) is mainly dependent on the distance and path-loss coefficient. All other types of noise are assumed to be negligible for the whole network. We assume omnidirectional antennas for both the BS and MS, and further assume that the effects of the uncovered regions due to circular geometry are negligible.

The analysis is based on the assumption that there are negligible errors when we approximate from hexagonal cell to a circular cell. Also, there may be nodes existing outside region-B (in order to make the interference calculations, for nodes close to the boundary, more complete). Hence, the formulas for interference power are consistent.

III. Analysis of the Hybrid Network Capacity

The purpose of the following analytical work is to find the optimum overall capacity for the HN, and then to find the optimum range of distance from the BS to place relay nodes. To achieve this purpose, we start with the calculation of interference level for both the ad hoc and the cellular networks. Circular geometry is used for calculations. Initially, the total interference for a receiving MS in ad hoc mode is determined and then SIR is computed for this link. This also enables us to calculate the capacity for this ad hoc link, C_{adhoc}. For the capacity of a cellular link, $C_{cellular}$, a straightforward methodology is adopted but with a constraint (to be described in Section III.B). When both C_{adhoc} and $C_{cellular}$ are obtained, the overall capacity of HN can be found. This also helps towards the optimal placement of the relay node.

A. Capacity Experienced by a Receiving Mobile Station (in Ad Hoc Mode) from a Relay Node

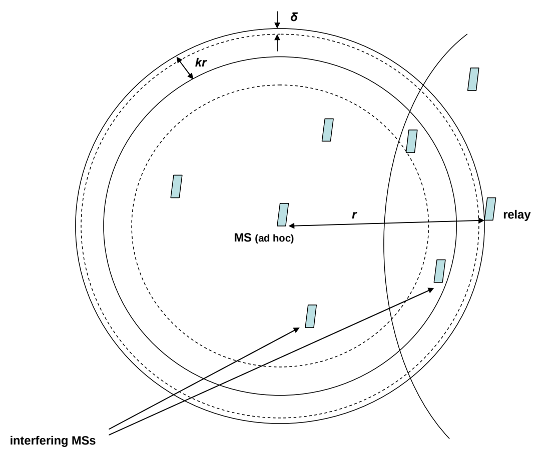


Figure A.2: Basic model for total interference calculation (ad hoc to relay).

Fig. A.2 is used as a basic model for the calculation of total interference experienced by a receiving MS from all MSs within the interference range from the MS to the relay. In this figure, r is the radius of the interference region of the MS (not BS, since here we are calculating only for the ad hoc part). [This r can be varied to see the effects on the capacity (to be described in Section IV).] This region is subdivided into $1/k$ sub-regions, where $0 < k << 1$. To perform continuous analysis, we consider, within a sub-region, a concentric annular ring with width δ between its outer and inner boundaries. The area of such a ring is given as

$$A_\delta = \pi r^2 - \pi (r-\delta)^2, \tag{1}$$

which is the elemental area for the whole HN.

Thus, any two MSs in the above mentioned ring may be considered to be at the same distance from the receiving MS at the center of the circle. The interference generated in this ring towards the receiving MS becomes

$$I_\delta = \frac{P_t \rho A_\delta}{x^\alpha}, \tag{2}$$

where x is the distance of an MS inside this ring from the receiving MS, ρ is the normalized node density, P_t is the transmit power of each MS, and α is the path-loss coefficient.

In each sub-region with width kr, there are sufficiently large number of rings each with width δ. Therefore, the integration of Eq. (2) with respect to δ from 0 to kr gives the interference contribution by a sub-region, I_{kr}, as

$$I_{kr} = \int_0^{kr} I_\delta d\delta. \tag{3}$$

At any instant, for the receiving MS, there is only one transmitting relay and all other transmitting MSs are simply contributing as interferers. The contribution of each of these interferers is also a function of their individual distance from the receiving MS. Hence the total interference, I_{total}, towards the receiving MS is obtained by integrating all the individual interference contributions over all the interference region of the receiving MS, (here, 0 to r). Thus,

$$I_{total} = \int_0^r I_{kr} dx = P_t \rho \left(\pi k^2 - \frac{\pi}{3} k^3\right) \left(\frac{r^{4-\alpha}}{4-\alpha}\right) \tag{4}$$

where P_t is the same for all MSs, and is fixed for a certain communication instant.

Now we consider the signal received at the MS. Given the distance between the relay node and the MS as r, the signal strength received at the MS, S_a, is obtained by

$$S_a = \frac{P_t}{r^\alpha}. \tag{5}$$

Since I_{total} is already found for a link between the sender and the receiver MS, the ad hoc signal-to-interference ratio, SIR_{adhoc}, can be calculated as

$$SIR_{adhoc} = \frac{S_a}{I_{total}} = K_1 \frac{4-\alpha}{r^4}, \tag{6}$$

where

$$K_1 = \frac{1}{\rho\left(\pi k^2 - \frac{\pi}{3} k^3\right)}. \tag{7}$$

Given SIR_{adhoc}, the ad hoc capacity can be found using the famous Shannon formula [1],

$$C_{adhoc} = B \log_2(1 + SIR_{adhoc}), \tag{8}$$

where B is the bandwidth of the channel.

B. Capacity Between the Relay Station and the Base Station

Considering Fig. A.3, the maximum possible distance for an MS from the BS is D. Given that a relay node is placed somewhere between the MS and the BS on a straight line, D can be simply divided into two parts, i.e., r, which is the distance between MS and relay node, and $D - r$, which is the distance between relay node and BS. Increasing r means placing a relay node closer to the BS, and vice-versa.

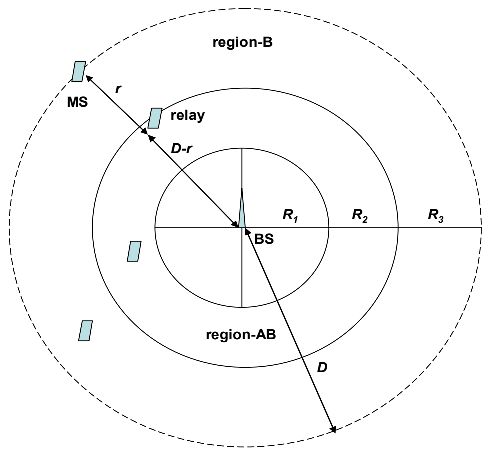


Figure A.3: Variation of r and placement of relay node.

According to [7], the SIR for a receiver in a cellular network (e.g., a CDMA system) can be approximated as

$$SIR = \frac{1}{N-1}, \tag{9}$$

where N is the number of MSs in the interference region of a cellular node. As we are using the continuous-analysis approach, the signal-to-interference ratio would also depend on the node density, ρ, and the distance, $D-r$. Since all interfering MSs are at most $D-r$ distance away from the BS, the number of MSs inside this range becomes $N = \rho\pi(D-r)^2$. Thus, the signal-to-interference ratio for the cellular link, $SIR_{cellular}$, is given by

$$SIR_{cellular} = \frac{1}{\pi\rho(D-r)^2 - 1}, \tag{10}$$

with a constraint

$$\pi\rho(D-r)^2 \geq 2, \text{that is, } r \leq D - \sqrt{\frac{2}{\pi\rho}}. \tag{11}$$

The above mentioned inequality (11) is imposed by considering that there should be at least two stations in this region in order to have the communication to take place. That is, one station is the transmitter (the relay node in this case) and the other station is acting as a receiver (the BS in this case). Other stations are regarded as interferers. The capacity for the cellular link, $C_{cellular}$, can now be calculated as

$$C_{cellular} = B\log_2(1 + SIR_{cellular}). \tag{12}$$

C. Total Capacity of the Hybrid Network

For the ad hoc network part, the capacity C_{adhoc} will decrease as r is increased (clear from Eqs. (6) and (8)). For the cellular network part, the capacity $C_{cellular}$ will increase as an MS moves closer to the BS (clear from Eqs. (10) and (12)). Thus, there would be a point in the HN where the overall total capacity of the HN is optimum. This means that there exists a cross point where the curves for C_{adhoc} and $C_{cellular}$ intersect, and this would be the total optimum capacity, C_{total}^{opt}, for the whole HN. Hence, the task of placing the relay node in an optimal location can be summarized in the following expression:

$$\begin{aligned} &\textit{maximize} && C_{total}^{opt} \\ &\textit{subject to} && d_{MS-relay} + d_{relay-BS} = D, \end{aligned} \tag{13}$$

where $d_{MS-relay}$ is the distance between the MS and its relay node and $d_{relay-BS}$ is the distance between the relay node and the BS. In other words, $C_{total}^{opt} = \max \{\min\{C_{adhoc}, C_{cellular}\}\}$. That is, we need to find an optimal value of r which gives maximum value for C_{total}^{opt}. This means to find an appropriate location for a relay node between an MS and the BS such that the maximum overall capacity for the whole HN is achieved.

Since the BS knows the number of MSs operating in a specific mode, it can dictate a node to work as a relay if it is within a certain range where optimum capacity is achieved. This is done by the BS by measuring the values of node density and path-loss constant.

IV. Numerical Results and Discussions

In this section we present the numerical results achieved from the analytical model of our scheme. MATLAB is used to analyze the performance of the system. An ideal medium access protocol is assumed. As mentioned earlier, the MSs are placed uniformly in the whole region of interest. The maximum interference region centered from the BS is a circle of radius $R_1 + R_2 + R_3 = D$. The maximum distance between the BS and the boundary of the network, D, is set as one kilometer. For illustration convenience, we assign this maximum radius at a normalized scale of $D = 10$ in Figs. A.4 and A.5. For example, an MS placed at distance $r = 7$ means that it is placed at a distance of 300 meters from the BS, of the whole possible distance of 1000 meters. Point 10 on the horizontal axis in all the graphs represents the location of the BS and point 0 represents the location of an MS at the boundary of the HN. It is again mentioned that there are some MSs assumed to be existing outside the boundary of the network in order to make the interference calculations complete. The vertical axis in all graphs illustrates the calculated capacity, C, in Mbps. The path-loss coefficient, α, is kept as 2 for Fig. A.4 and as 3.5 for Fig. A.5. The whole network region is divided into 100 concentric annular sub regions, i.e., $k = 0.01$.

Figs. A.4 and A.5 illustrate the capacity curves for the cellular link as well as the ad hoc link with respect to the distance from the BS. The dashed curves represent the ad hoc link capacity and solid curves represent the cellular link capacity. It can be easily observed that although the BS can be reached by nodes in region-B in case of pure cellular network, the capacity goes down as the distance from the BS increases. However, if we introduce a hybrid network, higher system capacity can be achieved. That is, when a relay node is placed between the BS and region-B, a better overall capacity may be achieved. By placing the relay node in most appropriate location, optimal network capacity can be reached, as described in the following two subsections.

A. Case I: Free-space Environment

Considering Fig. A.4, as expected, we observe that as r increases, C_{adhoc} decreases while $C_{cellular}$ increases. Thus, there is an optimal location, r_{opt}, for optimum overall capacity of the HN, which is obtained at the cross point of these two curves. Fig. A.4(a) depicts a sparsely populated network ($\rho = 0.2$). Here we achieve an optimum capacity of 3.87 Mbps at $r_{opt} = 7.4$. In Fig. A.4(b), which represents a moderately populated network, the optimum capacity decreases to 2.25 Mbps and is achieved at $r_{opt} = 7.9$. This is about 40% decrease with respect to the optimum capacity in Fig. A.4(a). Fig. A.4(c) represents a densely populated network. Here, the cross point of the capacity curves is achieved at $r_{opt} = 8$, with $C_{total}^{opt} = 1.52$ Mbps. Hence, the optimum capacity has dropped about 33% with respect to Fig. A.4(b) and about 60% with respect to Fig. A.4(a). This is due to the fact that as ρ increases, interference increases and thus the capacity decreases. Therefore, one needs to put a relay node closer to the BS in order to achieve the optimal capacity because more nodes are now generating interference. From the above observations, we conclude that an optimum range of distance from the BS for a relay node placement is found, which is, in this case, $r_{opt} \approx (7.4 \sim 8)$.

B. Case II: Urban Area Semi-open Environment

Fig. A.5(a, b, c) represents the results for a greater value of α, i.e., $\alpha = 3.5$, for the sparsely, moderately, and densely populated HN, respectively.

As shown in the figure, when we move the relay node closer to the BS, i.e., as r is increased, a change in optimum capacity is observed. However, the comparison of Fig. A.4(a) and Fig. A.5(a) shows that the cross point of the capacity curves in

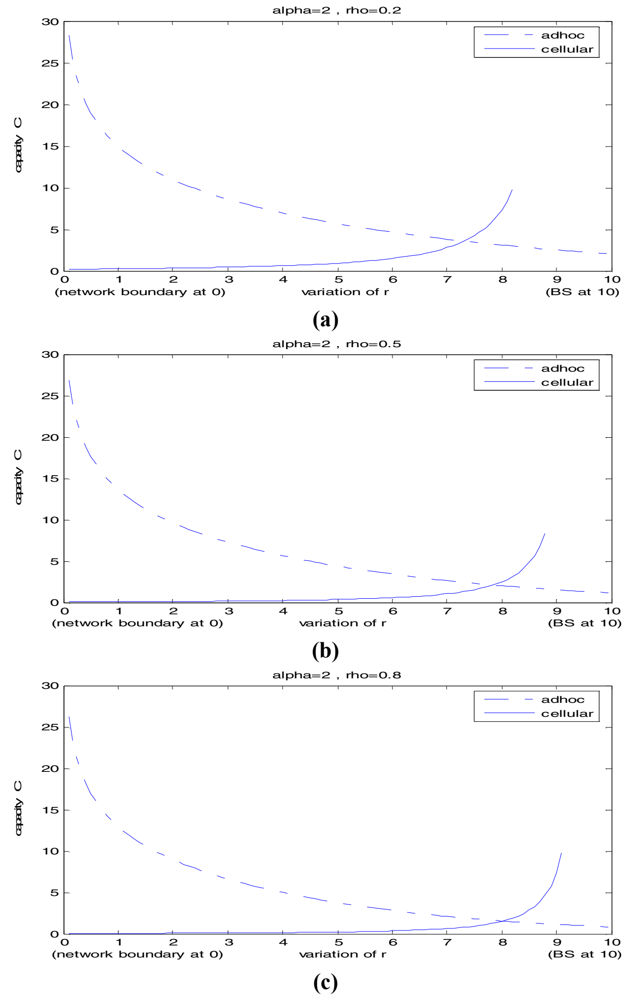


Figure A.4: Capacity curves for the hybrid network with lower path-loss constant: (a) Sparsely populated, (b) Moderately populated, (c) Densely populated.

Fig. A.5(a), (as $r_{opt} = 6.8$), is lower than the cross point in Fig. A.4(a), The reason for this is that as α increases, the capacity decreases for both network components (provided all the other relevant parameters are kept the same). This means, in order to achieve optimal overall capacity, one needs to place the relay node closer to the BS. Similarly, Fig. A.4(b) can be compared with Fig. A.5(b), and Fig. A.4(c) with Fig. A.5(c), and the same effect can be observed, i.e., a corresponding drop in the capacity with increase in α. More specifically, the achieved C_{opt}^{total} in Fig. A.5(a,b,c) are 2.28, 1.14, and 0.8 Mbps, respectively. These values are substantially lower than the corresponding values we observed in the free-space environment. However, again, an optimum range for the relay node placement is found, which is $r_{opt} \approx (6.8 \sim 7.3)$.

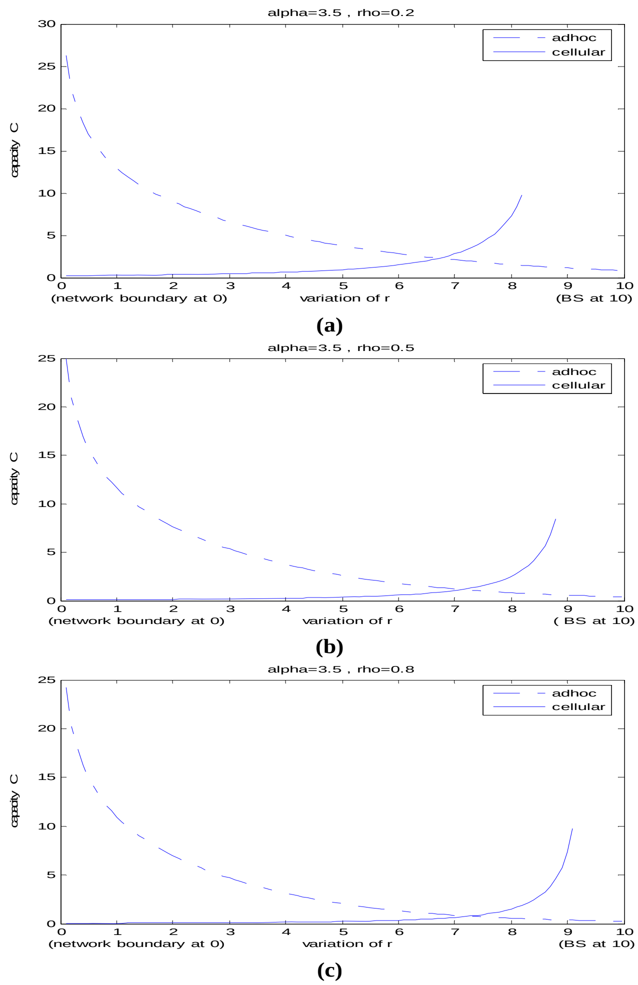


Figure A.5: Capacity curves for the hybrid network with higher path-loss constant: (a) Sparsely populated, (b) Moderately populated, (c) Densely populated.

From both Fig. A.4 and Fig. A.5, two important observations can be made. Firstly, the overall optimum capacity of the HN does not improve if a relay node is placed beyond a certain distance from the BS. However, a range for optimal placement of the relay node is found, as $r_{opt} \approx (6.8 \sim 8)$, in general. Secondly, for known values of α and ρ, a precise location for the relay node can be found which ensures maximum overall capacity for the HN.

V. Conclusions and Future Work

In this paper, we have proposed an approach of finding the optimum range for relay node placement in a Hybrid Network consisting of an ad hoc component and a cellular component. The major contributions of this paper are that we have proposed a method which employs continuous-space circular geometry for analyzing the optimum overall capacity of HN and that we have obtained optimal range of positions for relay node placement. From the numerical results, it is also found that, for known values of node density and path-loss coefficient, a *precise location* for a relay node placement can be determined which can give maximum overall capacity of the hybrid network.

As part of the future work, we are investigating the methods for efficient selection of a relay node when multiple relay nodes are available within the optimum range. Another possible future direction is to explore methods to improve the overall optimum capacity for such a scenario.

REFERENCES

[1] C. E. Shannon, A Mathematical Theory of Communications, *Bell Systems Technical Journal*, Vol. 27, 1948, pp. 379-423.

[2] I. Chlamtac, M. Conti, and J. J. N. Liu, Mobile Ad Hoc Networking: Imperatives and Challenges, *Elsevier, Ad Hoc Networks*, Vol. 1, No. 1, July 2003, pp. 13-64.

[3] D. Calvacanti, D. Agarwal, C. Cordeiro, B. Xie, and A. Kumar, Issues in Integrating Cellular Networks, WLANs and MANETs: A Futuristic Heterogeneous Wireless Network, *IEEE Wireless Communications*, Vol. 12, No. 3, 2005, pp. 30-41.

[4] H. Wu, C. Qiao, S. De, and O. Tonguz, Integrated Cellular and Ad Hoc Relaying Systems: iCAR, *IEEE JSAC*, Vol. 19, No. 10, 2001, pp. 2105-2115.

[5] L. K. Law, S. V. Krishnamurthy, and M. Faloutsos, Capacity of Hybrid Cellular-Ad hoc Data Networks, in *Proc. IEEE INFOCOM*, April, 2008, pp. 1606-1614.

[6] M. F. M. Hossain, A. Mammela, and H. Chowdhury, Impact of Mobile Relays on Throughput and Delays in Multihop Cellular Networks, in *Proc. ICWMC*, July, 2008, pp. 304-308.

[7] T. S. Rappaport, *Wireless Communications Principals and Practices*, 2nd ed, Prentice Hall, Inc., 2002.

Paper B

Analysis of Mobile-oriented Energy Consumption for Heterogeneous Connections in Hybrid Wireless Networks

---

Title: Analysis of Mobile-oriented Energy Consumption for Heterogeneous Connections in Hybrid Wireless Networks

Authors: Ziaul Haq Abbas and Frank Y. Li

Affiliation: Department of Information and Communication Technology, University of Agder (UiA), N-4898 Grimstad, Norway

Journal: *International Journal of Communication Networks and Distributed Systems (IJCNDS)*, vol. 9, no. 3/4, pp. 184–204, 2012. Available at http://www.inderscience.com/info/jhome.php?jcode=ijcnds

---

Analysis of Mobile-oriented Energy Consumption for Heterogeneous Connections in Hybrid Wireless Networks

Ziaul Haq Abbas and Frank Y. Li
Department of Information and Communication Technology
University of Agder (UiA), N-4898 Grimstad, Norway
E-mails: {ziaul.h.abbas, frank.li}@uia.no

Abstract — **While more and more mobile devices are equipped with both cellular and WLAN interfaces for Internet access, the energy consumption aspect of these connections has not been studied in depth in the literature. In this paper, a hybrid network with three alternative wireless connections, namely, a cellular component, a combined component with a mixed ad hoc and cellular link, or an infrastructure-based WLAN component, is investigated. A distance-based analysis on the energy consumption of a mobile station has been performed for each alternative wireless connection, and corresponding numerical results have been obtained. The results suggest that, to increase link throughput while maintaining minimal power consumption, the remaining energy of the mobile station as well as the amount of bytes communicated over a link should be taken into consideration for connection selection. Furthermore, the remaining battery level of the mobile station could also be used as the criterion for handover decision, from both the mobile station's and the network operator's perspectives.**
Keywords — hybrid wireless networks; energy consumption analysis; relay station; ad hoc; WLAN; cellular; battery lifetime.

I. Introduction

Internet access has become ubiquitous in recent years through not only wired but also wireless links. Most wireless-enabled devices such as laptops or Mobile Stations (MSs) are being used to access the Internet services via either cellular systems or wireless networks. Alongside, more and more MSs are being equipped with heterogeneous communication interfaces to support, e.g., Wireless Local Area Network (WLAN) and cellular connections. Generally speaking, the achieved data rate in the cellular mode is lower than that in the WLAN mode while the communication ranges of these two networks behave reversely.

One of the most restrictive constraints in designing new features for mobile devices is the limitation on battery power [11], [13], and [7]. Only small improvements in battery capacity are expected in the near future [11], predicted by the projections on the progress in the battery technology. On the other hand, extensive multimedia applications are storming into mobile devices, introducing increasingly higher demands for battery power. Under such circumstances, it becomes imperative for mobile industry to utilize battery efficiently. Thus a careful analysis of the energy consumption of the MS, no matter it is operating in an infrastructured network associated with a Base Station (BS)/an Access Point (AP), or in an ad hoc network without infrastructure support, becomes important. Furthermore, efficient utilization of energy may provide not only positive gain in the overall operation time of the MS itself but also implicitly contributes towards longer network connectivity of the whole network.

Given specific channel condition in wireless communications, for a fixed transmission power level, the supported data rate may be traded off with the distance between the MS and its one-hop communication counterpart, e.g., BS/AP or another MS which acts as a Relay Station (RS) to forward the MS's data. To transfer packets over long distance, the idea of sending data packets in a multi-hop fashion towards the destination by relaying the packets through intermediate nodes (i.e., RSs) has been shown of being able to reduce required transmission power [10]. Furthermore, it is also interesting to investigate which RS to select as the next hop towards the BS if more than one candidate RSs are present within the communication range of the sending MS. For RS selection in such a scenario, the energy consumption of the RS battery also needs to be considered in order to maximize the operation time of the selected link while delivering maximum amount of data. To investigate the energy consumption of an MS/RS, it is necessary to consider the power costs of transmitting, receiving, and processing a packet. Moreover, the relationship between energy consumption at a given transmission data rate for a certain distance between transceivers and the amount of information bytes been communicated (and thus the battery operation time) is important and needs to be explored.

In this paper, the energy consumption analysis is carried out for an MS which is part of a Hybrid Network (HN) consisting of a cellular component, a WLAN component with AP, and an ad hoc component. The MS may access the Internet through one of the three alternative links, i.e., (a) a direct cellular link via the BS, (b) a direct WLAN link through the AP, or (c) a combined link via an RS towards the BS. To perform the analysis, the energy consumed by the MS is computed for each of the above mentioned links. Moreover, the energy consumption of the RS is also calculated. Based on such energy information, the MS may handover between these possible links in order to optimize the overall operation time and the amount of the communicated data, as a tradeoff with the remaining battery energy. Furthermore, when the hybrid link is selected, the energy consumption information of the RS may also be utilized by the MS to decide which RS to select, provided that more than one RSs are available within the communication range between the MS and the BS. The major contributions of this study are twofold:

- A method for analyzing the remaining battery energy of the MSs for various distances between transceivers over diverse links for the purpose of optimal utilization of available resources has been developed.
- The calculated energy consumption results could serve as the basis for designing *mobile-centric* energy-aware handover and relay selection algorithms in IEEE 802.21-targeted HNs, as such energy-aware algorithms are not specified in the 802.21 standard [9].

The rest of the paper is organized as follows. Section II briefly summarizes the related work. In Sec. III, the scenario of the studied HN is described along with a few basic assumptions. Sec. IV describes the mechanisms for energy consumption analysis of the MS in different communication states for various links. Then, the numerical results are presented and discussed in Sec. V. More results with reference to RS as well as asymmetric Internet traffic are given in Sec. VI. Furthermore, some concrete directions for energy-aware hybrid link and relay selection, and handover strategy are pointed out in Sec. VII. Finally, the conclusions are drawn in Sec. VIII.

II. Related Work

The interest of the research community in the field of HNs has grown in the last decade. Some work has also been carried out on developing handover policies for such HNs. However, little attention has been paid on the MS's energy consumption based on the amount of traffic in a selected network. In the following paragraphs, we summarize the related work done by some other researchers in the field of HNs, as well as the difference between their work and ours.

The IEEE 802.21 [9], an upcoming standard for HNs, proposes a framework to efficiently implement solutions for inter-technology seamless handover. However, it does not specify any algorithms or schemes to be considered for such handover process, neither any parameters for such a decision. In this study, we consider energy consumption as a major parameter for path selection and handover from both MS's and operator's perspectives.

Handover performance between Third Generation (3G) mobile systems and WLAN access networks was investigated through simulations in [14]. In their work, the authors used WLAN signal level thresholds as the handover criteria by assuming that the MSs support the IEEE 802.21 cross layer architecture. As a conclusion, the authors are in favor of using WLAN over Universal Mobile Telecommunication System (UMTS) given that the economic cost and data rate are the major factors in their consideration. However, the energy consumption of the MS is not considered as a decision-making factor in their work.

Moreover, the authors in [2] developed an experimental setup to implement certain parts of an IEEE 802.21-based framework. They concluded that by tolerating longer delays and higher packet loss to an acceptable level, the IEEE 802.21-assisted handover helps in performing reliable proactive handover. However, how energy consumption may affect a handover decision is not investigated in their work either.

In [6], the author proposed a handoff algorithm, based on energy consumption measurements of UMTS and IEEE 802.11 WLAN networks specifically on an Android mobile phone, using the estimated application traffic load. This work is relatively experimental and focuses solely on the application layer of the communication protocol. Nevertheless, the effects of distance between transceivers on the corresponding battery energy consumption are not considered in this work. Moreover, only one of these two networks is utilized for uplink and downlink traffic before and after the handoff.

In [16], a model was set up by the authors to calculate handover cost in a predictive IEEE 802.21-oriented network architecture. However, no energy consumption effects on the MS are considered in their model.

In our earlier work [1], we considered an HN consisting only of a cellular and an ad hoc component. However, in this paper, we have also involved the WLAN link in the overall HN. Furthermore, a more thorough analysis has been carried out for the MS/RS energy consumption.

III. Heterogeneous Networks: Scenario and Assumptions

This section describes the architecture of the HN under consideration. Fig. B.1 gives the pictorial representation of the studied scenario. As shown in the figure, the overall HN is divided into three major parts, i.e., a WLAN component, a cellular component, and a hybrid component including an ad hoc link and a cellular link.

In the WLAN component, the MS can access the Internet via an AP at High Data Rate (HDR), and this link is referred to as a *WLAN link*. In the cellular component,

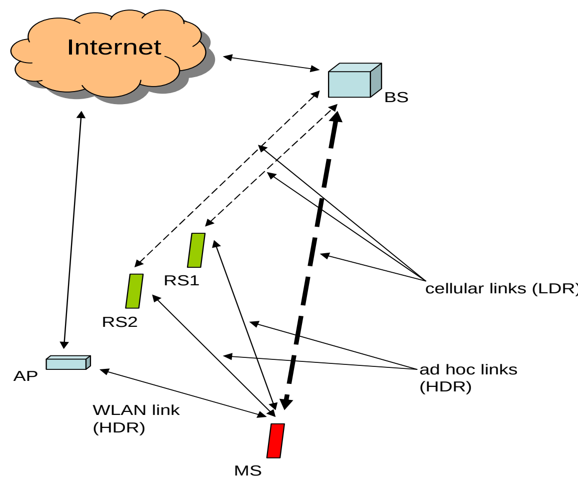


Figure B.1: Basic heterogeneous network architecture.

the MS can access the Internet at Low Data Rate (LDR) through a BS directly. This link is referred to as a *cellular link*. In the combined component, the MS uses an RS, which is another MS supporting both the cellular mode and the ad hoc mode, to reach the BS for accessing the Internet, in a two-hop fashion. Therefore, this two-hop link is referred to as a *hybrid link* and the link between the MS and the RS is referred to as an *ad hoc link*. To support the hybrid link, the MS(s) and the RS(s) are equipped with two radio interfaces, one for LDR cellular link and the other for HDR ad hoc/WLAN link.

There are various possible connections for the MS to establish communication with the Internet for both uploading and downloading. The connections considered in our work are (1) direct *cellular connection* in which the MS directly reaches the BS to connect to the Internet for both uplink and downlink; (2) *WLAN connection* in which the MS uses the AP to reach the Internet for both ways; (3) Uplink Relayed and Downlink Relayed *(URDR) connection* in which the MS uses an intermediate RS to reach the BS for both directions in a two-hop fashion, i.e., ad hoc and cellular; (4) Uplink Cellular and Downlink WLAN *(UCDW) connection* where the MS uses cellular link for uploading and WLAN link for downloading; (5) Uplink WLAN and Downlink Cellular *(UWDC) connection* where the traffic direction is opposite to that in the UCDW connection; (6) Uplink Relayed and Downlink WLAN *(URDW) connection* in which the MS uses an intermediate RS to reach the BS for the uplink and receives directly from the AP for the downlink; and (7) Uplink WLAN and Downlink Relayed *(UWDR) connection* in which the traffic flow is opposite to that in the URDW connection. Other heterogeneous combinations such as Uplink Cellular and Downlink Relayed (UCDR) and Uplink Relayed and

Downlink Cellular (URDC) may exist in the network however are not included here since the constituent individual link energy consumption for these paths is already implicitly reflected in the other paths described in the paper. For example, the energy consumption of the RS in UCDR and URDC is already reflected in the UWDR and URDR connections respectively. Another type of possible connection is that the MS may connect to the gateway via multiple relays in a multi-hop fashion. This is however beyond the scope of this paper.

The AP and the BS in the architecture are assumed to be powered by Alternating Current (AC) power lines and are hence without energy constraint, while the MS and the RSs are handheld devices with limited battery energy. The BS and the AP reach the Internet using wired connections.

Moreover, we assume that all mobile stations support both the Point Coordination Function (PCF) and the Distributed Coordination Function (DCF) [8], for radio channel access in the WLAN link and the ad hoc link, respectively. The power consumed for switching between the reception and transmission states as well as between the cellular link and the ad hoc/WLAN link is assumed to be negligible. Furthermore, the power consumption of the MS in the process of association/disassociation with the AP is also ignored. Idle-mode power consumption is neglected for the MSs/RSs.

For calculation convenience, we consider only one transmission cycle. The RS has sufficient buffer size to store the incoming data before forwarding it to the next station. The propagation delay is negligible. The control and *DATA* frames will not be lost and are received as error-free packets. A station cannot receive and transmit simultaneously. Furthermore, hidden terminals and capture effects are ignored.

The analysis is based on the assumption that only the AP can poll the MSs (because for an MS, receiving *Beacon* frames consumes much less power than periodically transmitting polling frames in an active manner). The MSs do not miss any *Beacons*, i.e., they always wake up immediately before each *Beacon* arrival. The contention-free parameter set elements [3] are included in the *Beacon* frame to keep all MSs apprised of contention free operation duration. All the MSs receiving this *Beacon* set their Network Allocation Vector (NAV) to the maximum duration of Contention Free Period (CFP) to block any DCF-based access to the medium. Moreover, the Contention Period (CP) does not overrun the CFP and is long enough for the transfer of at least one maximum-size *DATA* frame and its associated control and acknowledgement frames.

The effects of distance between the transceivers on their corresponding energy consumption are taken into account for all the above mentioned connections/links.

IV. Power Consumption Analysis

In this section, we analyze the power consumption of the MS/RS for the links described in Sec. III. As the first step, the calculations for power consumption in both the receiving and the transmitting states for an MS in the cellular mode are carried out. A simplified path-loss model [5] is utilized for this computation along with a condition placed on the minimum power received at the receiver if another MS is also transmitting to the same receiver simultaneously. This leads to the average power consumed per communicated byte which is taken as the average of power consumed per byte for both transmission and reception of a data packet. Later on, the average power consumption per byte for the MS operating in the ad hoc mode is computed. The power consumption analysis in both the cellular and the ad hoc modes enables us to calculate the power consumption of the RS in the hybrid link. In the end, the average power consumption per byte is calculated for the MS operating in the WLAN link (with AP). The expressions for the power consumption in the cellular, ad hoc, and WLAN links further enable us to calculate power consumptions for the various mixed connections mentioned in Sec. III. Furthermore, conventional energy-power relation is used to calculate the corresponding values for energy consumption. These results provide for the basis for designing handover strategies and for appropriate RS placement/selection in the considered HN, as discussed in Sec. VII.

A. Power Consumption in the Cellular Link

The analysis for the cellular link is carried out by considering one or more MSs within the coverage of a BS with Code Division Multiple Access (CDMA) connections. Note that our cellular link energy model is not generalized for any type of cellular systems, but targeted only at CDMA-based networks. The power consumed per bit by the MS, P_c, can be divided into two parts [12], i.e., the power needed in order to generate P_t amount of transmission power and the electronic processing power of the transmitter/receiver circuitry used mainly for encoding/decoding the transmitted/received signal. [It is worth mentioning that P_c is the power *consumed by the MS*, not the transmission power, in order to generate P_t amount of transmission power; and $P_c \neq P_t$.] Hence, we can write

$$P_c = k_1(P_{et} + P_t) + k_2(P_{er}), \tag{1}$$

where P_t is the transmission power, P_{et} and P_{er} are the electronic processing power for transmission and reception, respectively. The other two parameters in Eq. (1), k_1 and k_2, are binary integers reflecting the status of the communicating radio, respectively. Given that the MS radio cannot transmit and receive simultaneously, we have $k_1 = 1$ and $k_2 = 0$ when the MS is in the transmission state and $k_1 = 0$ and $k_2 = 1$ when it is in the reception state.

Since more circuitry is required to receive and decode the signal as a usual case, we relate P_{et} and P_{er} as $P_{er} = 2.5P_{et}$. [As suggested in [15], P_{er} is usually 2 to 3 times higher than P_{et}.] Hence, P_c can be approximated as

$$P_c = k_1(P_{et} + P_t) + k_2(2.5P_{et}). \tag{2}$$

Now, as indicated in Fig. B.1, assuming that another MS (e.g., an RS) is also transmitting at the same time, we need to vary the transmission power of the MSs in order to satisfy the fundamental requirement in CDMA that the difference between the received individual signal levels at the BS is within 1 dB [4]. Mobile stations need to adjust their transmission power accordingly in order to minimize the near-far effects at the BS receiver, and this procedure is controlled by the BS. Therefore, we can write

$$\left| \frac{X_1 P_{t_1}}{d_1^{\alpha}} - \frac{X_2 P_{t_2}}{d_2^{\alpha}} \right| \leq 1dB, \tag{3}$$

where P_{t_1} is the transmission power of the concerned MS, P_{t_2} is the transmission power of the other MS/RS, d_1 and d_2 are their corresponding distances to the BS, respectively, and α is the path-loss coefficient. For Eq. (3), the relationship between the received power and the distance is taken from the simplified path-loss model as $P_r = AP_t/d^{\alpha}$, where P_t and P_r are the transmitted and received power, respectively, and d is the distance between the transmitter and the receiver. X_1, X_2, and A are the unitless constants depending on corresponding antenna characteristics and channel conditions. Due to the consideration that distance is a major factor for path loss in our model and for the sake of expression simplicity, the effects from other factors like shadowing and multipath fading are not included in our formulae.

The average energy consumption, E_c, of the considered MS per communicated byte in the cellular link is simply calculated as

$$E_c = 8 \times P_c/R_c, \tag{4}$$

where R_c is the supported cellular data rate.

The transmission power of the signal must be taken into account because it needs to be adjusted with respect to the distance of transmitting MS/RS from the BS and the channel conditions in order to satisfy Eq. (3). Hence, the MS is required to adjust its P_t which also influences the corresponding energy consumption of the battery. The variation of the energy consumption per byte for the MS with respect to the varying distance from the BS, when another MS/RS is also transmitting at a power level of 300 mW from a constant distance of 600 meters from the BS, is shown in Fig. B.2.

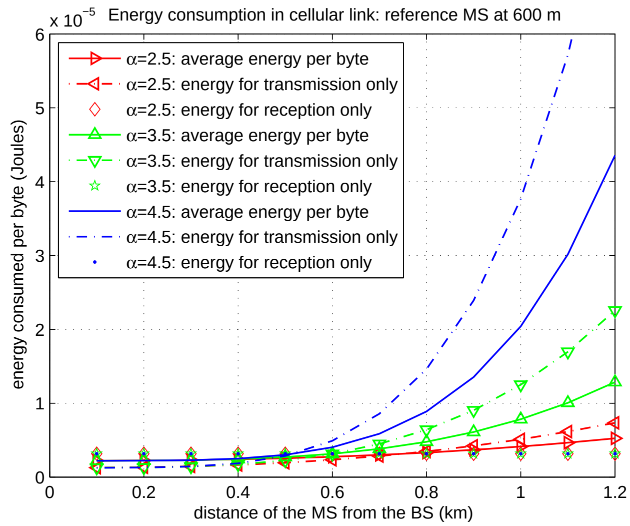


Figure B.2: Cellular link energy consumption: effects of α and distance.

As obvious, it is evident from Fig. B.2 that higher values of α (e.g., worse channel conditions) and longer distance from the BS increase the requirement on P_t of the MS to reach the BS and this correspondingly increases the energy consumption by the MS. Similarly, the energy consumption of the RS in cellular link is also affected by α and the distance from the BS.

One interesting observation from Fig. B.2 is that the energy consumption for reception-only is higher than both the energy consumption for transmission-only and the average energy consumption when the MS is closer to the BS. The reason for this effect is due to a smaller contribution of α and the distance parameter d to the average energy consumption when the MS is closer to the BS and channel conditions are moderate. This observation may potentially be utilized as a hint on using the MS in the cellular mode for uploading the data when it is closer to the BS.

B. Power Consumption in the Hybrid Link

For a communication channel, the capacity, C, in bits per second, can be calculated by the Shannon's law

$$C = W\ log_2(1+SNR) = W\ log_2(1+\frac{P_r}{P_N}), \quad (5)$$

where W is the signal bandwidth in Hertz and SNR is the signal-to-noise ratio of the received signal power P_r to the noise power P_N including all noise sources.

From Eq. (5), we can obtain the minimum P_r for a specific data rate, which when put in the expression for the simplified path-loss model, $P_r = XP_t/r^{\alpha}$, gives the minimum transmission power, P_t, needed to achieve the rate C at a distance r between the transceivers. [This is the raw data rate. The effects of different modulation and coding schemes as well as protocol overhead are not counted.]

Now, we firstly develop equations for MS power consumption in the ad hoc link. The expressions for power consumption for the hybrid link can then be obtained by using these ad hoc link equations and the cellular link equations (developed in Section IV.A). All stations are assumed to support PCF which runs on top of the DCF. Access to the medium is controlled by PCF during CFP and by DCF during CP, as indicated by Fig. B.3. The AP can dictate the duration of Contention Free Repetition Interval (CFRI) (and thus, implicitly, the duration of CFP and CP) [3]. For the ad hoc link, stations operate in the CP and use DCF. Here we calculate the power consumption of an MS/RS in a CP.

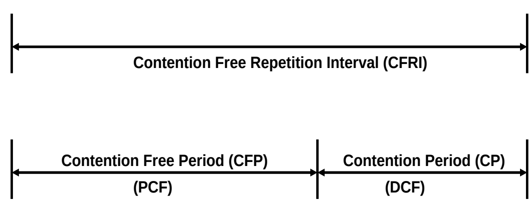


Figure B.3: Division of CFRI into CFP and CP.

During the CP, the interaction between the transmitter and the receiver follows the Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) mechanism, as shown in Fig. B.4. In this figure, the solid amplitude levels of a frame imply that the main power contribution in this frame is due to transmission. Similarly, the dotted amplitude levels imply the power consumed during reception. For a detailed description of the abbreviations RTS, CTS, DIFS, SIFS, and ACK, please refer to [8].

Now using Fig. B.4, the *total power* consumed by an MS to transmit the *DATA* frame and to transmit/receive the associated control frames during CP using DCF becomes

$$\begin{aligned} P_{tr\_total\_CP(DCF)} &= \{T_{DIFS}P_{DIFS}^{E} + E[T_{bo}]P_{bo} + T_{RTS}P_{RTS}^{T+E} + 3T_{SIFS}P_{SIFS}^{E} \\ &+ T_{CTS}P_{CTS}^{R+E} + T_{DATA}P_{DATA}^{T+E} + T_{ACK}P_{ACK}^{R+E}\}/T_1, \end{aligned} \quad (6)$$

where E, R, and T in the superscripts of the variable P indicate the power consumed in processing, receiving, and transmitting phases of the corresponding frames respectively. The subscript *bo* implies *backoff*. A variable T, with a subscript, rep-

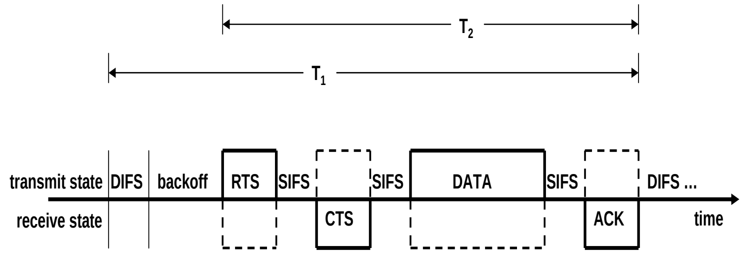


Figure B.4: Interaction between transmitter and receiver during CP.

resents the time consumed by a corresponding frame. The total time for one cycle, T_1, taken by the *DATA* and associated control frames is given by

$$T_1 \quad = \quad T_{DIFS} + E[T_{bo}] + T_{RTS} + 3T_{SIFS} + T_{CTS} + T_{DATA} + T_{ACK}. \tag{7}$$

Similarly, the total power consumed by the other MS for the reception of the *DATA* frame and transmission/reception of its associated control frames during CP using DCF becomes

$$\begin{aligned} P_{re\_total\_CP(DCF)} \quad &= \quad \{T_{RTS}P^{R+E}_{RTS} + 3T_{SIFS}P^{E}_{SIFS} + T_{CTS}P^{T+E}_{CTS} \\ &+ \quad T_{DATA}P^{R+E}_{DATA} + T_{ACK}P^{T+E}_{ACK}\}/T_2, \end{aligned} \tag{8}$$

where T_2, which is the total time taken by the *DATA* frame and the associated control frames, is given by

$$T_2 = T_{RTS} + 3T_{SIFS} + T_{CTS} + T_{DATA} + T_{ACK}. \tag{9}$$

Eqs. (6) and (8), and the relation $Energy = Power \times Time$ can now be utilized to find the corresponding energy consumptions by the MS for the transmission and reception of the *DATA* frame and the associated control frames.

An RS, while forwarding the data on the uplink, receives from the MS using the ad hoc link interface and transmits towards the BS using the cellular link interface. Similarly, on the downlink, the RS receives from the BS using the cellular link interface and transmits towards the MS using the ad hoc link interface. The average energy per byte consumed by the RS is the sum of average per-byte energy consumed in these two-hop links. This has also been taken as the energy consumed per byte by the RS in the hybrid link.

C. Power Consumption in the WLAN Link

There are two types of access periods that the MS may utilize in the HDR links, i.e., CP and CFP. For the ad hoc link, the MS uses CP and its analysis has already been presented in Section IV.B. For the WLAN link, the MS operates in CFP utilizing the PCF access rules for communication. An example of the kinds of frames used inside CFP [3] is shown in Fig. B.5. A *Beacon* frame announces the start of the CFP and the total duration of CFRI. [Just one *Beacon* frame with its corresponding CFRI is used for the numerical calculations.] The standard $DATA + CF\_ACK + CF\_POLL$ frame is used by the AP in the WLAN network during CFP, where *DATA* and CF_POLL are intended for the same MS and CF_ACK is intended for the previous transmission. *SIFS* is used to separate these frames. CFP may be foreshortened by PCF at or before the CFP maximum duration based on the size of the polling list and available traffic. A CF_END frame is transmitted by the AP to announce the end of a CFP. Contention period follows CF_END.

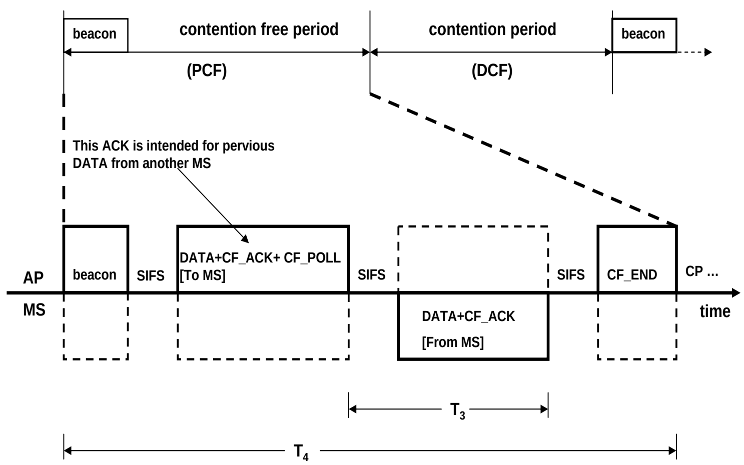


Figure B.5: Subdivision of CFP.

During CFP, the total power consumed by the MS in the transmission mode for the *DATA* frame and the associated control frames can be expressed as

$$P_{tr\_total\_CFP(PCF)} = \{T_{SIFS}P^{E}_{SIFS} + T_{DATA+CF\_ACK}P^{T+E}_{DATA+CF\_ACK}\}/T_3, \qquad (10)$$

where T_3, the total time taken by the *DATA* frame and the associated control frames in this case, is given by

$$T_3 = T_{SIFS} + T_{DATA+CF\_ACK}. \qquad (11)$$

An MS, after receiving the *Beacon* frame, remains active during the whole CFP. Hence, it is always in the active state [3] during the CFP. The corresponding power consumption in the reception mode for the *DATA* frame and the associated control frames can be accordingly expressed as

$$\begin{aligned} P_{re\_total\_CFP(PCF)} &= \{T_{beacon}P^{R+E}_{beacon} + 3T_{SIFS}P^{E}_{SIFS} + T_{CF\_END}P^{R+E}_{CF\_END} \\ &+ P^{R+E}(T_{DATA+CF\_ACK+CF\_POLL})\}/T_4, \end{aligned} \tag{12}$$

where T_4, which is the total time taken by the *DATA* frame and the associated control frames, is given by

$$T_4 = T_{beacon} + 3T_{SIFS} + T_{DATA+CF\_ACK+CF\_POLL} + T_{CF\_END}. \tag{13}$$

Correspondingly, the energy consumption of the MS in the WLAN link can be found by using Eqs. (10) and (12), and the relation $Energy = Power \times Time$.

D. Power Consumption in the Mixed Connections

The power consumption relations developed for the cellular, ad hoc, and WLAN links in Secs. IV.A - IV.C, respectively, can now be used to calculate the power consumption of the MS (and RS) for the various connections described briefly in Sec. III. In the rest of this section, the subscript *pb* denotes *per byte*. The remaining subscripts are self-explanatory.

D.1. URDR Connection

The MS transmits to and receives from the RS on the ad hoc link in the URDR connection. The corresponding total power consumption per byte, $P_{tot\_URDR\_MS}$ can thus be written as

$$P_{tot\_URDR\_MS} = \{P_{tr\_total\_CP(DCF)}\}_{pb} + \{P_{re\_total\_CP(DCF)}\}_{pb}. \tag{14}$$

The RS receives on the ad hoc link from the MS and transmits on the cellular link for the uplink traffic. The traffic path is in the opposite direction for the RS in the downlink. The total average power consumption per byte for the RS, $P_{tot\_URDR\_RS}$ using the URDR connection can thus be expressed as

$$\begin{aligned} P_{tot\_URDR\_RS} &= \{P_{re\_total\_CP(DCF)}\}_{pb} + \{P_{et} + P_t\}_{pb} + \{P_{er}\}_{pb} \\ &+ \{P_{tr\_total\_CP(DCF)}\}_{pb}. \end{aligned} \tag{15}$$

The total two-hop link power consumption can now be written as

$$P_{tot\_link\_URDR} = P_{tot\_URDR\_MS} + P_{tot\_URDR\_RS}. \tag{16}$$

D.2. UCDW and UWDC Connections

In the UCDW connection, the MS transmits directly on the cellular link and receives from the AP on the WLAN link. The traffic direction in the UWDC connection is opposite to that in the UCDW connection. The corresponding total power consumption for these connections can be respectively given as

$$P_{tot\_UCDW\_MS} = \{P_{re\_total\_CFP(PCF)}\}_{pb} + \{P_{et} + P_t\}_{pb}, \tag{17}$$

$$P_{tot\_UWDC\_MS} = \{P_{tr\_total\_CFP(PCF)}\}_{pb} + \{P_{er}\}_{pb}. \tag{18}$$

D.3. URDW and UWDR Connections

In the URDW connection, the MS transmits through the ad hoc link to the BS and receives from the AP on the WLAN link. The traffic direction in the UWDR connection is opposite to that in the URDW connection. The corresponding power consumption equations can be written respectively as

$$P_{tot\_URDW\_MS} = \{P_{tr\_total\_CP(DCF)}\}_{pb} + \{P_{re\_total\_CFP(PCF)}\}_{pb}, \tag{19}$$

$$P_{tot\_UWDR\_MS} = \{P_{tr\_total\_CFP(PCF)}\}_{pb} + \{P_{re\_total\_CP(DCF)}\}_{pb}. \tag{20}$$

The RS in the URDW connection receives on the ad hoc link from the MS and transmits on the cellular link to the BS; while in the UWDR connection, it receives from the BS and transmits to the MS. The corresponding power consumption per byte in these connections can be respectively given as

$$P_{tot\_URDW\_RS} = \{P_{re\_total\_CP(DCF)}\}_{pb} + \{P_{et} + P_t\}_{pb}, \tag{21}$$

$$P_{tot\_UWDR\_RS} = \{P_{er}\} + \{P_{tr\_total\_CP(DCF)}\}_{pb}. \tag{22}$$

The total two-hop link power consumption for the URDW and UWDR connections respectively becomes

$$P_{tot\_link\_URDW} = P_{tot\_URDW\_MS} + P_{tot\_URDW\_RS}, \tag{23}$$

$$P_{tot\_link\_UWDR} = P_{tot\_UWDR\_MS} + P_{tot\_UWDR\_RS}. \tag{24}$$

The corresponding energy consumption values can be again obtained using the conventional relation between energy and power.

It is worth mentioning that Eqs. (16), (23), and (24) have not been explicitly used in the numerical results presented later in this paper. However. they are presented here for the sake of completeness.

V. Numerical Results for MS Energy Consumption with Symmetric Traffic

In this section we present the numerical results achieved from the above analysis considering symmetric traffic between the MS and the Internet. MATLAB is used to evaluate the performance of the system with different parameter configurations. CSMA/CA has been used as the underline MAC protocol for energy consumption calculation during CP and a fixed length of 300 bytes has been assumed for the *DATA* frame. For the WLAN link, the energy consumed by the MS is calculated by considering the MS as operating in the CFP only. PCF decides about the duration of CP and CFP, and the association or disassociation of the MS with the WLAN link. For the cellular link, a data rate of $R_c = 2$ Mbps is assumed. For the HDR ad hoc and WLAN links, the data rate of $R_a = R_w = 54$ Mbps is used. The values of the path-loss coefficient are taken as $\alpha_c = 2.5$ for the cellular link; while for the ad hoc and WLAN links, they are taken as $\alpha_a = \alpha_w = 4$ respectively. Representative distance values are selected as $d = 200, 400$, and 600 meters for the cellular link; $r_a = 100$ and 150 meters for the ad hoc link; and $r_w = 100, 150$, and 200 meters for the WLAN link, respectively. Extra protocol overhead is neglected during the numerical computations. Table B.1 summarizes configurable parameters [12] used for the analyses of the ad hoc and WLAN links.

Table B.1: Ad hoc and WLAN link analysis parameters.

Parameter	Value	Parameter	Value
basic rate	6 Mbps	T_{CTS}	14*8/basic rate
data rate	54 Mbps	T_{SIFS}	10 μs
DATA length	300 bytes	T_{DIFS}	50 μs
P^E	0.635×10^{-7} W	T_{RTS}	20*8/basic rate
$E[T_{bo}]$	80 μs	T_{beacon}	179*8/basic rate
		T_{ACK}	14*8/data rate

Fig. B.6(a,b) through Fig. B.11(a,b) illustrate respectively the total bytes communicated and the total battery operation time for an MS (and/or RS) in different links/connections with respect to the remaining energy in the battery. The initial individual value of battery energy is taken as 10 Joules for all the mobile devices. In the following, we discuss, and briefly compare, the options that the MS can utilize for communicating the data to and from the Internet.

A. Direct Cellular Link

As shown in Fig. B.6(a), for the same amount of consumed energy, the number of total bytes exchanged by the MS while using direct cellular link is much higher if it is placed closer to the BS (i.e., lower values of d). The reason for this result is that the MS requires more power to transmit the signal when it is farther away from the BS. This causes more energy consumption from the MS and the battery depletes quicker with larger d. Similarly, the reason for larger separation of the curves at higher values of d and α is the relation between transmission power and (α, d). The higher the value of α and d, the more the power required by the MS and hence the quicker the battery energy depletion. This is also evident from the nonlinear relation between energy and (α, d) in Fig. B.2.

(a) (b)

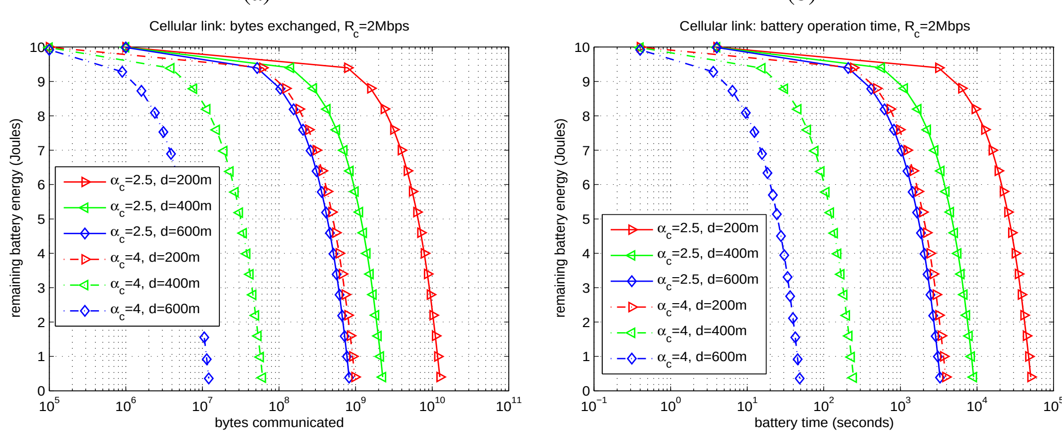


Figure B.6: Direct cellular link: relation between remaining energy of MS and (a) bytes exchanged (b) battery lifetime.

By comparing Fig. B.6(a) and Fig. B.6(b), similar trends can be observed for the total communicated data and the total battery operation time with respect to (α, d). The higher the values of α and d, the lower the total operation time. The reason for this result is that at higher values of α and d, more average energy is consumed per byte by the MS (due to higher required transmission power) and hence the battery depletes quicker, leading to correspondingly shorter battery operation time.

B. WLAN Link with AP

In the WLAN link (and in the ad hoc link), the data rate has been kept as constant. For this reason, the transmission power has to be increased if the distance between the transceivers is larger. This would also increase the energy consumed

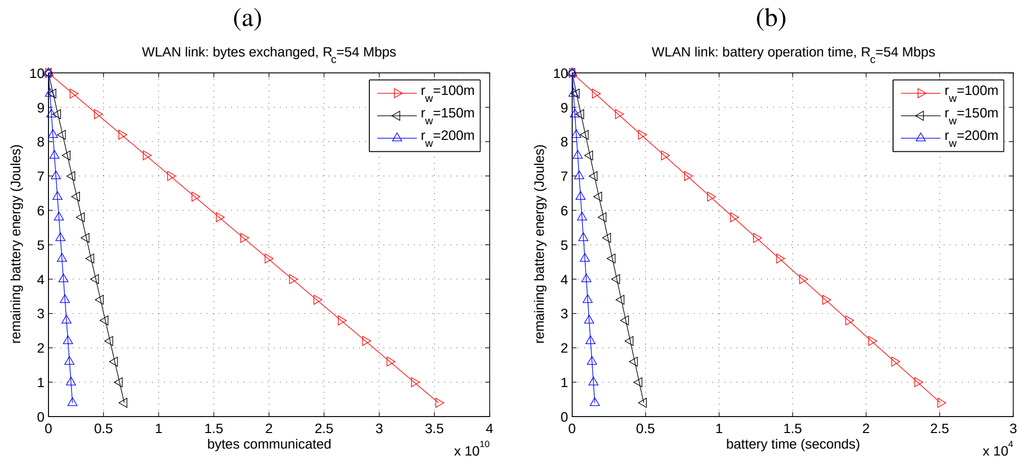


Figure B.7: WLAN link: relation between remaining energy of MS and (a) bytes exchanged (b) battery lifetime.

per byte for the MS/RS.

Fig. B.7 shows the number of bytes exchanged and the battery time for the MS operating in the WLAN link. As evident from the figure, for the same level of consumed energy and the same data rate, with larger values of r_w (α_w kept constant), a lower amount of data is exchanged. However, with just 100 meters of increase/decrease in the distance between the MS and the AP, there is a large amount of corresponding decrease/increase in the number of exchanged bytes. The reason for this effect is that with a higher value of data rate used in the WLAN link, the amount of transmission power used per byte increases/decreases rapidly with corresponding increase/decrease in the distance between the transceivers. Fig. B.7(b) is also consistent with Fig. B.7(a). The higher the value of r_w, the more the transmission power required, and hence, the lower the total battery operation time.

An interesting observation by comparing Figs. B.6(a) and B.7(a) is that for $d =$ 200 meters (keeping $\alpha_c = 2.5$), the MS exchanges more bytes with the BS than the number of bytes exchanged between MS and AP in the WLAN link when $r_w = 200$ meters. However, as hinted in the above paragraph, with decrease in r_w, the MS in the WLAN gains more in terms of exchanged date. Thus, *when the distance of the MS from the BS and the AP is comparable, moving the MS closer to the AP will gain more in terms of exchanged bytes.* However, this gain is obtained at the expense of certain degree of loss in the battery time (evident from the comparison of Figs. B.6(b) and B.7(b)).

C. Hybrid URDR Connection

The relationship of energy consumption with the amount of data exchanged and the battery time is respectively shown in Fig. B.8(a and b) for both the MS and the

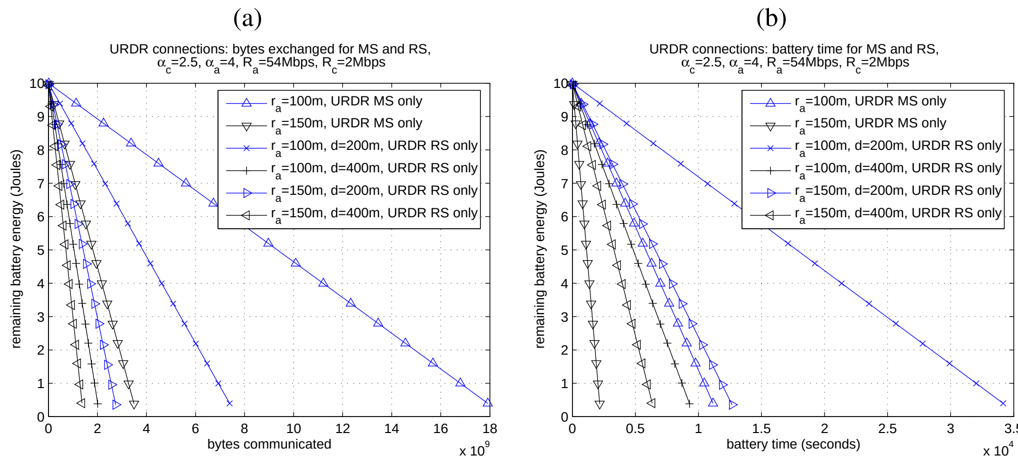


Figure B.8: URDR connection: relation between remaining energy of MS and RS (a) bytes exchanged (b) battery lifetime.

RS in the URDR connection. It is worth mentioning again that in this connection, the MS only uses ad hoc link while the RS uses the combined cellular and ad hoc links.

In Fig. B.8(a), the amount of bytes communicated by the MS when $r_a = 100$ meters is much larger than when $r_a = 150$ meters, for the same level of consumed energy. This is because of the higher transmission power requirement for higher values of r_a in the HDR ad hoc link. However, comparing Fig. B.8(a) with Fig. B.7(a), when $r_a = r_w = 100$ meters, for the same level of consumed energy the MS exchanges more data in the WLAN link. The reason for this effect is that the MS has to contend for channel access in the ad hoc link and hence it loses precious energy for accessing the channel due to possible backoffs. Thus, on average, it consumes relatively more energy per byte in the ad hoc link, thereby, reducing the total number of exchanged bytes.

The effects of URDR connection on the RS will be described and compared later in Sec. VI.A when we discuss the URDW and UWDR connections.

D. Mixed Connections

There are two main categories of mixed connections described in this subsection.

D.1. UCDW and UWDC Connections

Fig. B.9 illustrates the relationship of energy consumption for the MS in the UCDW and UWDC connections for different values of d and r_w.

For a fair comparison, we consider only the plots in both connections for the same values of d and r_w, i.e., $d = 200$ meters and $r_w = 150$ meters for both the UCDW and the UWDC connections. As evident from Fig. B.9(a), for the same level

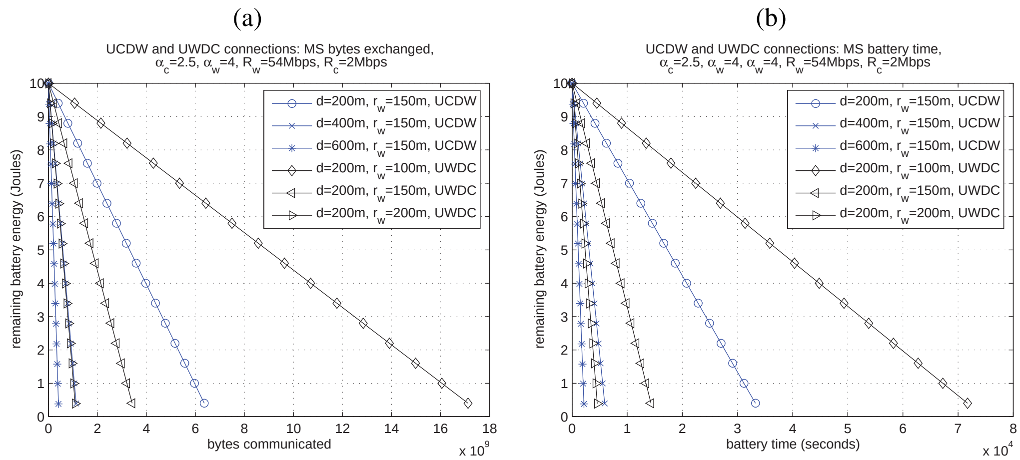


Figure B.9: UCDW and UWDC connections: relation between remaining energy of MS (a) bytes exchanged (b) battery lifetime

of energy consumption, the number of bytes exchanged by the MS in the UCDW connection is almost double the number in the UWDC connection (compare plots marked with ‘circle’ and ‘left-triangle’). The reason for this effect is that in the UWDC connection, the MS needs to transmit at a higher distance and with HDR and hence consumes more energy per byte. Thereby, the battery quickly depletes.

However, by comparing the plots marked with ‘cross’ and ‘right-triangle’ (i.e., comparing $d = 400$ meters and $r_w = 150$ meters in UCDW with $d = 200$ meters and $r_w = 200$ meters in UWDC), we observe that the amount of transferred bytes is almost the same. This gives us a crude conclusion that for reasonable and feasible distances, about 25% increase in the distance between the MS and the AP in the UWDC connection gives almost similar effects in terms of data transfer as an increase of more than 50% in the distance between the MS and the BS in the UCDW connection. Furthermore, when both cellular and WLAN connections are available, for shorter distances from the AP, *uploading through UWDC and downloading through UCDW leads to more gain* in terms of both the data amount and the battery lifetime.

D.2. URDW and UWDR Connections

A comparison of energy consumption for the MS in the URDW and UWDR connections is illustrated in Fig. B.10. For the MS only, this is effectively a comparison between the ad hoc and WLAN links (because it is only the RS which uses the cellular link in these mixed connections; and the discussion about RS is postponed till a later section to avoid complexity at this point).

Comparing the URDW and UWDR connections for $r_a = r_w = 100$ meters, we observe that, for the same amount of consumed energy, there is almost the same

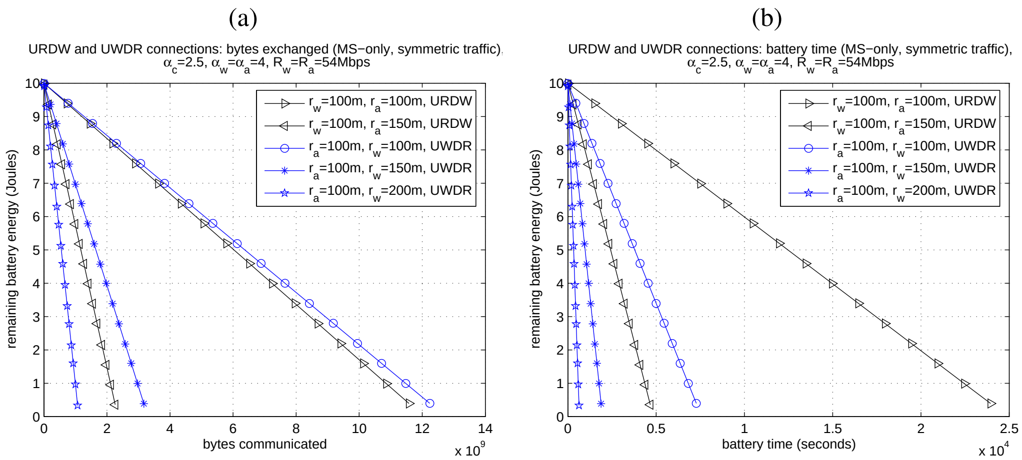


Figure B.10: URDW and UWDR connections: relation between remaining energy of MS (a) bytes exchanged (b) battery lifetime.

amount of data transferred for both types of connections. However, the battery time for these values of r_a and r_w is much more in the URDW connection than in the UWDR connection. The reason for this effect is that in the UWDR connection, the MS has to consume precious energy for receiving (and processing) a long *Beacon* frame which reduces total battery time. Furthermore, the MS has to transmit certain handshaking frames on the ad hoc link at the basic rate during the reception from the RS. Hence, *for comparable ad hoc and WLAN link distances, the URDW connection is preferred over the UWDR connection with respect to battery lifetime.*

To investigate the effects of increasing the ad hoc link distance and decreasing the WLAN link distance; or vice versa, compare the plots marked with 'left-triangle' and 'star' (i.e., $r_w = 100$ m, $r_a = 150$ m, URDW and $r_a = 100$ m, $r_w = 150$ m, UWDR), in Fig. B.10(a). For given similar ad hoc and WLAN link distances from the MS, moving the MS 50 m farther away from the AP has lower impact on the amount of total transferred data than moving the MS 50 m farther away from the RS in the ad hoc link. The effects on the corresponding battery times are also reversed and its reason has already been explained in the above paragraph.

VI. More Results for RS Energy Consumption and Asymmetric Traffic

In the previous section, our main emphasis was around the effects on energy consumption *of the MS* in various links/connections. However, in a mixed and/or relayed connection, more than one MSs (i.e., both MS and RS) are consuming energy simultaneously in order to exchange data for the MS. Hence, it is worth discussing the impact of different connections on the energy consumption of the RS as well.

(a) (b)

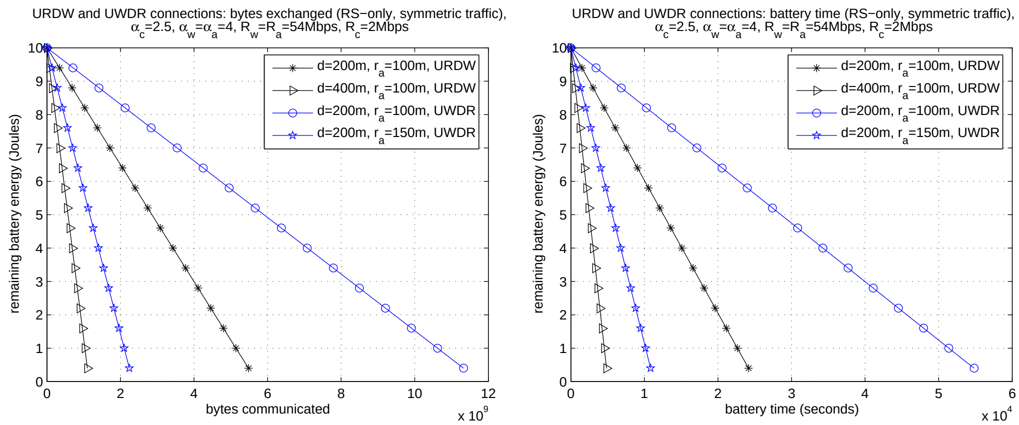


Figure B.11: URDW and UWDR connections: relation between remaining energy of RS (a) bytes exchanged (b) battery lifetime

Furthermore, asymmetric Internet traffic which is more realistic in practice is also considered in this section.

A. RS in the Mixed Connections

Fig. B.11 illustrates the energy consumption relation of the RS in the URDW and UWDR connections. Comparing these connections for $d = 200$ m and $r_a = 100$ m, we observe a clear advantage of using the UWDR connection in terms of both the amount of transferred data and the battery lifetime. The reason for this effect is that the RS has to transmit on shorter ad hoc links in the UWDR connection (it receives on the cellular link; and the energy consumed for reception is relatively little than the energy consumed for transmission) and hence the total average consumed energy per byte in this connection is lower than in the URDW connection. Therefore, *from RS's point of view, uploading through WLAN and downloading through relayed link enhances the battery lifetime while positively gaining in terms of the amount of transferred data.*

Furthermore, comparing Fig. B.11 with Fig. B.8 as an example, for $d = 200$ m and $r_a = 100$ m, the RS in the UWDR connection gains more in terms of both the amount of data exchanged and the battery time than the RS in the URDR connection. However, for these values of d and r_a, the RS in the URDR connection performs better than the RS in the URDW connection in terms of the amount of data transferred. Moreover, the battery lifetime for the RS for these values of d and r_a in the URDR connection is more than the corresponding RS lifetime in the URDW connection. Thus, *when all of these three types of mixed connections are available, for the same values of r_a and d, UWDR performs best in terms of the amount of transferred data.* This is because of the reason that in the UWDR connection, the RS transmits on the HDR ad hoc link at shorter distance consuming lower amount of energy. Moreover, the energy consumed by the RS for receiving on the LDR cellular link is very little, contributing towards a larger amount of data transfer.

B. Asymmetric Traffic: RS in URDW and UWDR Connections

Typically, hosts in the Internet generate relatively smaller amount of bytes for uploading than for downloading. For the sake of brevity, only the case of bytes exchanged by the RS in the URDW and UWDR connections is described, as illustrated in Fig. B.12. In this figure, it is assumed that 10% of the traffic is for uploading and 90% is for downloading.

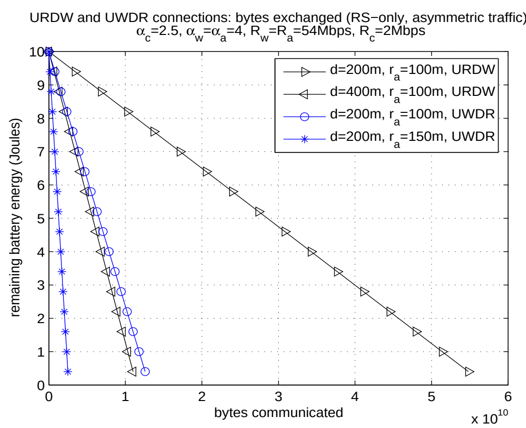


Figure B.12: URDW and UWDR connections: relation between remaining energy of RS and bytes exchanged for asymmetric traffic

Comparing the URDW and UWDR connections for fixed values of d and r_a (i.e., $d = 200$ m and $r_a = 100$ m), we observe (from the plots marked with 'right triangle' and 'circle', respectively) that for the same amount of consumed energy the RS in the URDW connection exchanges relatively more bytes. The reason for this effect is that the RS in the URDW connection receives (on the ad hoc link) and transmits (on the cellular link at a shorter distance) only 10% of the total traffic. In the UWDR connection, RS has to transmit 90% of the total traffic on the ad hoc link. This link consumes a major portion of the RS energy, resulting in a lower amount of total transferred bytes because of quicker battery energy depletion.

VII. Concrete Directions on Link Selection and Handover Strategies

This section briefly presents some concrete directions for developing link selection and handover algorithms in HNs according to the network topology as presented in Fig. B.1. While the results discussed in Sec. V and Sec. VI may be utilized mainly for developing mobile-centric handover strategies, they may also be used when developing handover algorithms from a network operator's perspective. Furthermore, considerations on the selection/placement of the RS, when the hybrid link is selected, are also presented.

Unlike the conventional handover definition, by *mobile-centric* handover, we assume that the MS may make its own decision to change its connectivity from one type of network to another type in an HN. This may be based on the information about the possible amount of data that may be transferred and the battery energy consumption (or battery lifetime) if a certain link is chosen. Similarly, the *network-instructed* handover may be performed between possible mixed connections based on the total information about battery lifetime and data quantity for each of the potential connections.

A. Mobile-centric Energy-aware Link Selection Strategies

Assume that the MS is aware of the current energy status of its own battery and that of the potential RSs within its ad hoc network. The MS may decide to handover, in an egotistic manner (assuming this is allowed by the network and the procedure is done in a co-ordinated fashion), between different links/connections based on how much advantage it would get by selecting one of the the given alternative links/connections. For example from Fig. B.7(b), considering that the MS is initially in the WLAN link, operating at 54 Mbps at a distance of 150 meters from the AP, the total operation time it can have in this link is about 83 minutes and the total bytes it can communicate is about 0.7×10^{10}. If it decides, based on an energy-aware link selection criterion, to switch to the direct cellular link, it would gain advantage both in terms of the communicated bytes and the operation time (considering $\alpha_c = 2.5$ and $d = 200$ meters in Fig. B.6(b)). However, this gain is obtained at the expense of slower data transfer rate. On the other hand, if the MS is located far away from the BS (e.g., $d = 600$ meters), there is almost no advantage of switching to the cellular link. It indicates that *if an MS is not far away from the BS and the channel conditions are moderate, it favors using the cellular link instead of the WLAN link, especially when battery lifetime is a major factor for decision making*. When the MS is far away from the BS, however, the hybrid link may help in terms of both energy conservation and lifetime extension.

As another example, suppose that there are two options for the MS: either to use the direct cellular link or to use the hybrid link, depending on channel conditions and application sensitivity (for example, delay tolerance). Under poor cellular channel conditions, the MS may decide to use the RS on the first-hop and then the RS may use the cellular link to reach the BS on behalf of the MS. This is a case of the URDR connection. However, using the hybrid link leads to an extra cost, i.e., both the MS and the RS are simultaneously losing energy. Hence, the MS gains advantage but at the expense of another MS's (i.e., the RS's) lifetime, which may affect the combined link lifetime.

B. Handover Strategies from Network's Perspective

The above subsection only mentions a few possibilities for handover decision merely from an MS's perspective. More comprehensive handover strategies can be developed from the network perspective, e.g., whether for a specific connection the battery lifetime has priority over the speed of data transfer when considering both the battery lifetime and the amount of communicated data, etc. The network operator may consider other relevant information such as the number of MSs as well as their traffic pattern and load in the network. It also needs to estimate the network connectivity time for certain selected link/connections for Quality of Service (QoS) provisioning. Thus, we have to consider multiple parameters in such scenarios, with different focus on one or more parameters, e.g., the remaining battery operation times of the involved MSs/RS, the amount of communicated data per unit time (traffic load), the distance of the MS from the RS, BS and AP, etc. Different mixed connections for uplink and downlink traffic can also be utilized.

Such a strategy may also include the decision of switching to a different data rate based on the same link. Furthermore, certain weights may be assigned to various parameters based on different priorities for each specific application scenario.

C. Strategies for Relay Station Selection

Given that the MS decides to use the hybrid link to access the Internet in the presence of multiple relays, a question of how to select the best relay arises. As already explained and illustrated in Figs. B.6, B.8, and B.11, the values of d and α play major roles in the energy consumption of an MS and RS. Thus, when the MS is within the communication range of more than one RSs, it needs to select an RS for which the distance from the MS and the BS is optimum based on the parameters like d, r_a, battery lifetime, etc. Always selecting the RS closest to the BS may not be an optimum solution. However, collectively optimizing the energy consumption and the battery lifetime for both the MS and the RS may provide us with better overall network connectivity.

VIII. Conclusions and Future Work

In this paper we have proposed an approach for calculating energy consumption of an MS equipped with heterogeneous interfaces in a wireless HN architecture. Using the proposed distance-based mobile-oriented energy consumption analysis, the battery operation time and the amount of data communicated in each of the constituent links of the HN can be calculated. Moreover, various possible mixed connections are analyzed and discussed for MS/RS energy consumption. The results obtained in this work suggest that there is no single connection which performs best in terms of both the battery lifetime as well as the amount of transferred data for various distances between the involved transceivers. Thus, for developing an optimum energy-aware handover strategy many parameters need to be taken into account. As part of our future work, a feasible energy-based handover methodology for heterogeneous wireless networks will be developed by jointly considering the energy consumption at the MSs/RSs and other parameters as well as available resources of the network.

REFERENCES

[1] Z. H. Abbas and F. Y. Li, “Power Consumption Analysis for Mobile Stations in Hybrid Relay-assisted Wireless Networks”, in *Proc. 5th IEEE ISWPC*, Modena, Italy, May 2010.

[2] A. Dutta, S. Das, D. Famolari,Y. Ohba, K. Taniuchi,T. Kodama, and H. Schulzrinne, “Seamless Handover across Heterogeneous Networks - An IEEE 802.21 Centric Approach”, in *Proc. WPMC 2005*, Aalborg, Denmark, September 2005.

[3] M. S. Gast, *802.11 Wireless Networks The Definitive Guide*, 2nd edition’, O’Rielly and Associates Inc., 2005.

[4] I. A. Glover and P. M. Grant, *Digital Communications*, 2nd edition, Pearson-Prentice Hall, 2004.

[5] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.

[6] P. Henrik, “Energy-Aware Network Selection using Traffic Estimation”, in *Proc. 1st ACM workshop on Mobile Internet through Cellular Networks*, Beijing, China, September, 2009.

[7] IEEE “Special Issue on Energy-aware Ad Hoc Wireless Networks”, *IEEE Wireless Communications*, vol. 9, issue. 4, August, 2002.

[8] IEEE *IEEE 802.11: LAN/MAN WIRELESS LANS*, IEEE Standards Association. Website http://standards.ieee.org/getieee802/802.11.html

[9] IEEE IEEE 802.21, Media Independent Handover Interoperability Working Group.
Homepage http://www.ieee802.org/21

[10] M. Kubisch and H. Karl, "Analyzing Energy Consumption in Wireless Networks by Relaying", *Tech. Rep., Technical University of Berlin*, Berlin, Germany, June 2001.
Website http://www.tkn.tu-berlin.de/publications/papers/frame.pdf

[11] K. Lahiri, A. Raghunathan, S. Dey, and D. Panigarhi, "Battery-driven System Design: A New Frontier in Low Power Design", in *Proc. 7th ASP Design and Automation Conference*, Bangalore, India, January 2002.

[12] C. Levy-Bencheton and G. Villemaud, "Power Consumption Optimization in Multi-mode Mobile Relay", in *Proc. 12th European Conference on Wireless Technology (EuWiT)*, Rome, Italy, September 2009.

[13] J. R. Lorch A. J. Smith, "Software Strategies for Portable Computer Energy Management", *IEEE Personal Communications*, vol. 5, issue. 3, June 1998.

[14] A. de la Oliva, T. Malia, A. Vidal, C. J. Bernardos, I. Soto, and A. Banchs, "A Case Study: IEEE 802.21 Enabled Mobile Terminal for Optimized WLAN/3G Handovers", *Mobile Computing and Communication Review*, vol. 11, no. 2, April 2007.

[15] E. Shih, S. Cho, N. Ickes, R. Min, A. Sinha, A. Wang, and A. P. Chandrakasan, "Physical Layer Driven Algorithm and Protocol Design for Energy-Efficient Wireless Sensor Networks", in *Proc. ACM MOBICOM 2001*, Rome, Italy, July 2001.

[16] S. -J. Yoo, D. Cypher, and N. Golmie, "Timely Effective Handover Mechanism in Heterogeneous Wireless Networks", *Wireless Personal Comm*, vol. 52, no. 3, February 2010.

Paper C

Two Teletraffic-based Schemes for Energy Saving in Cellular Networks with Micro-cells

---

Title:	Two Teletraffic-based Schemes for Energy Saving in Cellular Networks with Micro-cells
Authors:	Ziaul Haq Abbas and Frank Y. Li
Affiliation:	Department of Information and Communication Technology, University of Agder (UiA), N-4898 Grimstad, Norway
Journal:	*Journal of Communications (JCM)*, Special Issue on Improving the Energy Efficiency of Cellular Communications, vol. 7, no. xx, pp. xx–xx, Accepted for Publication, 1st June 2012.
Copyright©:	Academy Publisher

Two Teletraffic-based Schemes for Energy Saving in Cellular Networks with Micro-cells

Ziaul Haq Abbas and Frank Y. Li
Department of Information and Communication Technology
University of Agder (UiA), N-4898 Grimstad, Norway
E-mails: {ziaul.h.abbas, frank.li}@uia.no

Abstract — **The energy consumption of Base Stations (BSs) is known to constitute a major part of the power consumption in a cellular network. In this paper, we propose a novel approach which may switch a BS off under light traffic conditions in order to conserve the power consumption of such networks. More specifically, when the traffic load in the middle cell of a network with three micro-cells is sufficiently low, the corresponding BS can be switched off and its users will be covered by increasing the transmission power of one sector antenna in each of the two neighboring cells. Two teletraffic-based power saving schemes are proposed in our study. The first scheme analyzes the expected sojourn times of different channel occupancies and switches off the BS *deterministically* when the switching thresholds are met. The second scheme instead switches off the BS *probabilistically* based on a policy designed using a Finite Markov Decision Process (FMDP). Numerical results for the first scheme demonstrate that a reasonable amount of network power can be saved at the cost of slightly higher transmission power. The results for the second scheme indicate that a lower limit on the long-term network transmission power can be obtained using the FMDP-based analysis.**

Index Terms — Micro-cell, teletraffic, power saving, energy, Markov chain, BS, FMDP, optimization.

I. Introduction

The ubiquity of the Internet and mobile networks nowadays is causing rapid growth in the number of cellular access networks, especially at the micro- and pico-cell levels. Due to this rapid growth, the energy consumption of such networks is becoming an important issue from both the environmental and the economical points of view. Furthermore, because of the ever increasing data services such as social networking and the huge jumps in the number of mobile phone subscribers, a larger quantity of infrastructure equipment is also required.

[0]Submitted 21st Oct. 2011, revised 15th Apr. 2012, accepted 1st Jun. 2012.
[0]The deterministic scheme presented in this work has been published as [12].

On the other hand, the cellular network operators feel an impetus towards reducing the energy consumption of their networks in order to decrease the expense of operating a network. Due to competition in the market, the operators try to reduce their network operation cost while still keeping high Quality of Service (QoS) for the customers. From a system-level point of view, more than two-thirds of the energy in a mobile communication network are consumed by the part responsible for radio access, mainly Base Stations (BSs). Therefore, the operators' focus is shifting from optimizing the network deployment schemes towards developing techniques, such as intelligently switching unnecessary BSs off in order to further reduce the energy consumption of already optimally deployed cellular networks.

Moreover, as the BSs in a telecommunication network consume a major portion of the total network power [1] - [3], the power consumption reduction of BSs may be viewed as a definite direction towards green communication at the system level. Furthermore, in the technological domain, operators have started to incorporate new features in their infrastructure equipment allowing network elements to be remotely controlled, and even switched off under certain circumstances [4] and [5]. Taking such advancements as motivation, we propose in this paper a network energy saving scheme which switches off the middle BS in a linear configuration of a network with three micro-cells when the traffic intensity in the corresponding cell is below a threshold level. Correspondingly, the users of the switched-off BS are covered by increasing the transmission power of a sector antenna from each of the two neighboring cells when their respective traffic intensities are sufficiently low as well.

More specifically, we propose two distinct schemes with different approaches on when and how to switch off the BS. With the first scheme, we analyze the transmission and total power for the network based on the expected sojourn times for different channel occupancies and switch the BS off or on deterministically when the border states are reached. With the second scheme, we intend to minimize the transmission power consumption of the network by switching off the BS with certain probability using a Finite Markov Decision Process (FMDP) based state transition policy.

In the rest of the paper, Sec. II describes related work while Sec. III presents the network scenario along with a few assumptions. Secs. IV and V respectively analyze the two studied schemes. Sec. VI gives an account of BS power consumption. Numerical results for the two studied schemes are presented and discussed respectively in Secs. VII and VIII. Sec. IX gives a brief comparison of the two schemes. Finally, the paper is concluded in Sec. X.

II. Related Work

The interest of the research community in the field of network energy conservation has grown in the past few years. A lot of work has been carried out on the energy consumption reduction of individual network components. However, little attention has been paid on the idea of reducing the network energy consumption by switching part of the network components off while ensuring service availability and continuity. In the following paragraphs, we summarize the related work done by other researchers.

In [6], the authors show that dynamic adjustment of cell size according to the significant traffic variation in a day can help conserve network energy. However, their solution may create blind spots in the geographical coverage of the network. Moreover, the authors of [7] evaluate energy saving that can be achieved using energy-aware cooperative management of two different cellular access networks offering services over the same geographical area. According to their proposal, when traffic intensity is so low that only one of these networks is sufficient to provide desired QoS for the users of both networks, the other network is switched off. However, this approach requires complex potential hardware modification in the involved stations as well as inter-network cooperation algorithms and inter-operator agreement.

Furthermore, in [5], a few concepts to save energy in small-cell wireless communication BSs are presented and one of them suggested to switch off parts of the BS components if traffic load is low. However, no detailed analysis of any of the presented concepts is given in [5]. The possibility to decrease energy consumption of a cellular network is investigated in [8] by reducing the number of active cells during the low-traffic periods according to the time in a day. To do so, the authors carry out an analysis for a 24 hour period by considering low traffic intensity especially in the early morning, late evening, and night times. However, in reality, traffic intensity may become high even during night times, e.g., during festivals, emergency situations etc. How their scheme would deal with such situations is not addressed by these authors. Therefore, a solution which is able to switch off (and on) cells (and/or sectors) according to the variation in traffic intensity, regardless of the time in a day, is needed. This observation indeed partially triggered our work in this paper.

Furthermore, an analytical model to determine the effectiveness of the policies that activate Access Points (APs) in dense Wireless LANs according to user demands is proposed in [9]. However, the authors do not present the optimality of the performance of the proposed policies in terms of energy savings. In [10], the authors propose an approach aiming to realize green radio system by introducing the deployment of Green Base Stations (GBSs) within the already deployed cellular network. GBSs are basically powered by solar and wind energy sources. Coverage optimization for energy saving is presented in their work by optimally increasing or decreasing the footprint of GBSs. However, this approach reduces energy consumption of the network of already existing BSs at the cost of infrastructure expenditure added by the deployment of GBSs. Furthermore, what is the optimal figure of the saved total network energy (considering both types of BSs) is not discussed in [10]. Moreover, the authors in [11] present an insight into optimizing the energy efficiency at the BS in a cell through power control, however, focusing only on the down-link data transfer.

III. Network Scenario: Power Saving Scheme and Assumptions

The network scenario studied in this paper is depicted in Fig. C.1(a). As shown

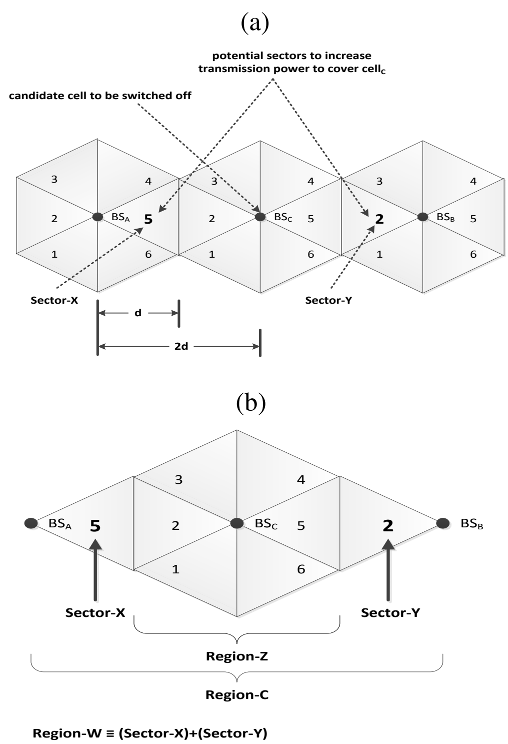


Figure C.1: Network scenario (a) Overall network; (b) Definitions of regions.

in the figure, there are three adjacent identical-size hexagonal cells in the network, namely $cell_A$, $cell_C$, and $cell_B$. Every cell has a corresponding BS, i.e., BS_A, BS_C, and BS_B. In the remainder of the paper, the terms $cell_j$ and BS_j would be interchangeably used, where $j = A, C$, and B. Every BS has six sector antennas to cover its corresponding geographical region. For illustration purpose, Sector 5 of $cell_A$ and Sector 2 of $cell_B$ are respectively denoted as *Sector-X* and *Sector-Y*. All the six sectors of $cell_C$ are collectively denoted as *Region-Z*, as also depicted in Fig. C.1(b). When a sufficient number of channels are available, Sectors X and Y are the potential sectors to increase the transmission power and cover the Mobile Stations (MSs) of $cell_C$ which will be switched off.

In our scheme, if there are enough channels available in Sectors X and Y and the total traffic intensity in Region-Z is sufficiently low then BS_A can increase the transmission power of the antenna for Sector-X to cover Sectors 1, 2, and 3 of Region-Z. Similarly, BS_B can cover Sectors 4, 5, and 6 of Region-Z by increasing the transmission power of its antenna for Sector-Y. Since all sectors of Region-Z are now supported by the neighboring two cells, $cell_C$ can be switched off. This mechanism can significantly reduce the total network power consumption while still providing coverage to all the users of the network (although only in low/moderate traffic situations). However, as Sectors X and Y have to increase their transmission power while the antennas of $cell_C$ are switched off (i.e., no transmission power consumption for $cell_C$), there would be a tradeoff between the total saved power and the extra consumed transmission power. How to minimize this transmission power consumption is one of the goals of this paper. It is worth mentioning that once a decision to increase the transmission power is taken, both Sectors X and Y increase their transmission power simultaneously and coordinately to fully cover Region-Z. Hence, from here on, the region collectively composed of Sectors X and Y would be referred to as *Region-W*, as depicted in Fig. C.1(b). Furthermore, the region comprising collectively of Regions W and Z is referred to as *Region-C*.

Moreover, we define two modes of power consumption for the system. When Region-W increases its transmission power and $cell_C$ is switched off, the system is referred to as working in the Power Save ON (PS-ON) mode. Otherwise, it is referred to as working in the Power Save OFF (PS-OFF) mode.

For the sake of analysis simplicity, we assume that the traffic intensity in Sectors X and Y is the same and that the traffic around the central vertical line in Region-C is symmetric. It is also assumed that the transmission power increments of a sector antenna, and of its corresponding MSs under coverage, remain within the transmission regulation limits. That is, when a sector antenna increases its transmission power to cover the users at a distance of $2d$, the transmission power still remains under the regulation limits. Therefore, the proposed scheme is not suitable for cellular networks with macro-cells. A simplified path-loss model is utilized between the MSs and the BSs. Handshake procedures for cooperation among the BSs are

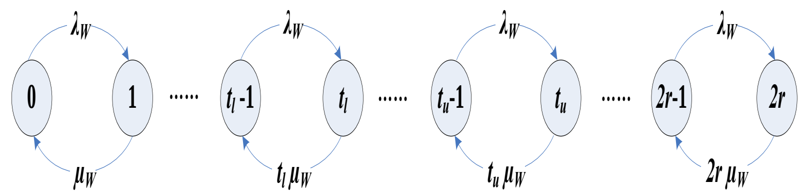


Figure C.2: BD process for Region-W.

ignored. For the power increasing of the antennas for Region-W, and the corresponding BS switching off, a controller is assumed to be operating on the backend. Furthermore, the switching latency that the BSs need for transmission status change is also ignored.

For calculation convenience, only homogeneous traffic with identical bit rate service is considered. Furthermore, for MSs and BSs, both uplink and downlink power control schemes are assumed to be operating [13]. When the system operates in the PS-ON mode, the interference effects around the vertical line dividing Region-Z into two equal halves (i.e., the line separating Sectors 1, 2, and 3 from Sectors 4, 5, and 6) are ignored. Moreover, the traffic around the mentioned vertical line in Region-C is also assumed symmetric.

IV. Deterministic Scheme Based on Sojourn Time Analysis

This section presents the power consumption analysis of our first scheme, i.e., sojourn time-based deterministic scheme.

Each of the regions W, Z, and C are individually modeled using separate Markov Chains (MCs). Each state in these MCs represents the collective number of occupied channels in the corresponding region. Thereafter, we obtain the blocking probabilities for each case from the MCs.

Based on our analysis, a power saving policy is developed with a few conditions being imposed on the individual thresholds of the number of channels occupied in each sector/region in order to avoid frequent mode switching between the PS-ON mode and the PS-OFF mode.

A. Markov Chains for the Regions

Each of the Regions W, Z, and C is modeled as a Birth-Death (BD) process, with Poisson arrival rate and exponential service times, using MCs. As an example, Fig. C.2 shows the process for Region-W which has $2r$ channels in total. In this figure, the arrival rate is denoted by λ_W and the departure rate by μ_W. The initial state of the process is 0 and the final state is $2r$, where r is also the maximum number of channels available in a sector.

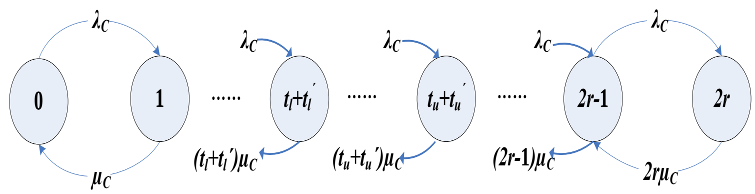


Figure C.3: BD process for Region-C.

For the BD process of Region-Z (whose MC is not shown as a figure due to its minor pictorial variation from Fig. C.2), the arrival and departure rates are represented as λ_Z and μ_Z, respectively. Since six sectors are considered collectively as Region-Z, there are a total number of $6r+1$ possible states for the MC of Region-Z. The lower and upper threshold values of channel occupancy in Regions W and Z, denoted respectively by t_l, t_u and t'_l, t'_u, would be discussed in Subsection IV.B.

However, as Region-C which jointly represents the area covered by Regions W and Z is obtained by increasing the transmission power of Region-W, there are still only $2r$ channels available in Region-C. The MC for Region-C is illustrated in Fig. C.3.

Each state of the MC for Region-C represents the sum of the channels occupied in Region-W and Region-Z. Hence the arrival and departure rates are respectively denoted by $\lambda_C \equiv \lambda_W + \lambda_Z$ (because independent arrivals are considered in Regions W and Z) and $\mu_C \equiv \mu_W$ (because, like Region-W, only $2r$ channels are available to Region-C, as it is only a geographical (and not channel wise) extension of Region-W). The threshold values in Region-W and Region-Z are used by Region-C to decide whether the system operates in the PS-ON or the PS-OFF mode.

The blocking probabilities for each sector can be obtained according to the Erlang-B formula. For example, for Region-W

$$p_{2r} = \frac{(\frac{\lambda_W}{\mu_W})^{2r}/(2r)!}{\sum_{j=0}^{2r}(\frac{\lambda_W}{\mu_W})^j/j!} \quad , \qquad j = 0,1,2,\ldots,2r. \tag{1}$$

Here, p_{2r} is the probability that all $2r$ channels are occupied.

The blocking probability can be taken as a Grade of Service (GoS) measure in a system. In a loss system, GoS is used to indicate the proportion of calls that are lost due to congestion. Thus for a particular required GoS, the traffic intensity offered by each user (i.e., λ/μ) and the number of channels in use, we can find the total offered traffic intensity through the Erlang table. This leads to the number of users that can be supported in a region [14]. For decision making about mode switching, the blocking probability can then be utilized as a criterion, i.e., the threshold values of channel occupancy can be adjusted according to the blocking probability requirements.

B. Power Saving Policy

Our power-saving policy is primarily based on the threshold channel occupancy values of Region-W and Region-Z, considered collectively in the MC for Region-C.

Figure C.4 illustrates the policy. As shown in the figure, when the MC is in any of the states from 0 to $t_l + t_l'$, the network operates in the PS-ON mode. Similarly, when the MC is in any of the states from $t_u + t_u'$ to $2r$, the network operates in the PS-OFF mode. However, when the system reaches state $t_l + t_l'$ or state $t_u + t_u'$, mode switching may happen.

To prevent the system from frequent (and undesirable) switching between the PS-ON and the PS-OFF modes, there is a hysteresis region defined between states $t_l + t_l'$ and $t_u + t_u'$. This means that if the system reaches state $t_u + t_u'$ from any of the states between $t_u + t_u' + 1$ and $2r$ (i.e., the number of occupied channels is decreasing and the system is already in the PS-OFF mode), it does not simply trigger the PS-OFF mode until it reaches state $t_l + t_l'$. On the other hand, if the system is in the PS-ON mode and the number of occupied channels is increasing, it will remain in the same mode till it reaches the other threshold $t_u + t_u'$. Thus, states $t_l + t_l'$ and $t_u + t_u'$ are triggering states. There is no decision-making within the hysteresis region, i.e., between $t_l + t_l'$ and $t_u + t_u'$.

There are a few necessary conditions/requirements for our model to be effective, as given below:

1) thresholds t_l and t_l' are respectively less than t_u and t_u';

2) $t_u + t_u' \leq c_1 \times 2r$ and $t_l + t_l' \geq c_2 \times 2r$. Here, c_1 and c_2 are scalers chosen according to the requirement/priority given to the PS-OFF and the PS-ON modes, respectively, and $0 < c_2 < c_1 < 1$. This guarantees that the upper threshold in Region-C is lower than $2r$ and the lower threshold is greater than

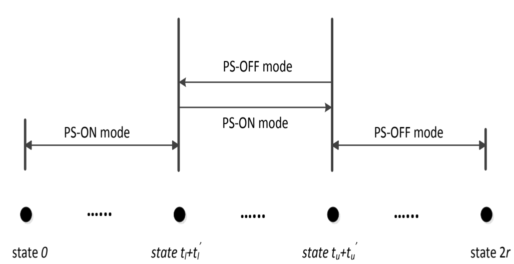


Figure C.4: Power saving policy for the network.

0, and there is a non-zero hysteresis region;

3) for the MCs of all sectors to be stable, $\lambda_i < \mu_i$, $i \in \{W, Z, C\}$.

*C. The Matrix **S***

To evaluate the possible benefit of energy saving, we need to calculate the expected amount of time the system stays in the PS-ON and the PS-OFF modes. For this purpose, from the MC of Region-C, the sum of the expected times that the process remains within lower threshold is calculated. Thus, we calculate the total expected time of the process in state $t_l + t_l'$ given that it starts from any of the states $0, 1, \ldots, t_l + t_l'$ (i.e., a pure PS-ON mode). Furthermore, according to the above described power saving policy, we also need to obtain the total expected time that the process remains in the PS-ON mode (the middle part of Fig. C.4) till it reaches state $t_u + t_u'$, i.e., the triggering state for the PS-OFF mode. Summing these two total times (i.e., the time in the pure PS-ON mode and the time in the hysteresis region while moving towards the upper threshold) gives the total time for the PS-ON mode, T_{PS-ON}.

Similarly, the total expected time the system spends in the PS-OFF mode, T_{PS-OFF}, can be calculated. The expected time within the hysteresis region is shared by both the PS-ON and the PS-OFF modes.

The above mentioned expected times can be obtained from a matrix **S** whose entries represent the expected sojourn times in different transient states, as described in the following.

From the BD process of Region-C, the rate transition matrix **Q** can be obtained. Using **Q**, the probability transition matrix **F** is computed[1] through the uniformization method[15] , i.e., each element q_{ij} of **Q** is transformed to f_{ij} of **F** as

$$f_{ij} = \begin{cases} q_{ij}/\theta & if \quad i \neq j, \\ 1 + q_{ii}/\theta & if \quad i = j, \end{cases} \tag{2}$$

with θ chosen such that $\theta \geq \max |q_{ii}|$. Please note that q_{ii} are the entries on the main diagonal of **Q** for the BD process and thus are negative numbers.

Let s_{ij} denote the expected number of time periods that the MC (for Region-C) is in state j given that it starts in state i. Let **S** denote the matrix of values s_{ij}, $i, j = 0, 1, \ldots, 2r$, i.e., $\mathbf{S} =$

[1] In **F**, a few of the row sums are less than 1 while others are 1. This is because that the elements of **F** are the probabilities from one *transient* state to another *transient* state.

$$\begin{vmatrix} s_{0,0} & \cdots & s_{o,t_l+t'_l} & \cdots & s_{o,t_u+t'_u} & \cdots & s_{0,r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{t_l+t'_l,0} & \cdots & s_{t_l+t'_l,t_l+t'_l} & \cdots & s_{t_l+t'_l,t_u+t'_u} & \cdots & s_{t_l+t'_l,r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{t_u+t'_u,0} & \cdots & s_{t_u+t'_u,t_l+t'_l} & \cdots & s_{t_u+t'_u,t_u+t'_u} & \cdots & s_{t_u+t'_u,r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{r,0} & \cdots & s_{r,t_l+t'_l} & \cdots & s_{r,t_u+t'_u} & \cdots & s_{r,r} \end{vmatrix}$$

The matrix **S** is easily obtained from **F** as [16]

$$\mathbf{S} = (\mathbf{I} - \mathbf{F})^{-1}, \tag{3}$$

where **I** is an identity matrix.

Now, using **S**, the mean times spent in the PS-ON and the PS-OFF modes (as mentioned above) can be calculated by summing the relevant elements of **S**.

Note that in our numerical results, to be presented later in Secs. VII and VIII, T_{PS-ON} and T_{PS-OFF} are normalized with the total time, T, of **S**, and thus are *unitless numbers*. Moreover, this implies that the energy consumption of these modes is indeed reflected even though the figures are labeled as power.

By using T_{PS-ON} and Eqs. (12) and (13) of Sec. VI, the total transmission and network power consumption in the PS-ON mode can be found. Similarly, T_{PS-OFF} can be used to calculate the power consumption values in the PS-OFF mode.

V. Probabilistic Scheme Based on FMDP Analysis

In the sojourn time-based scheme presented above, once the system reaches a triggering state, the mode is deterministically changed. In this section, we present the transmission power consumption analysis for our second scheme, i.e., FMDP-based scheme, in which the switching decision is made probabilistically instead. That is, the decisions to switch to the PS-ON or the PS-OFF mode are taken with certain probability (to be explained in the following).

Initially, we briefly present the necessary traffic analysis utilized for this scheme, followed by an example of how such analysis can be helpful in the policy development for FMDP. A state-transition matrix **P** is introduced, followed by the FMDP and the policy for minimizing the network transmission power consumption. Later on, a cost function, which is the long-term expected transmission power consumption, is defined and minimized using linear programming.

A. Traffic Analysis

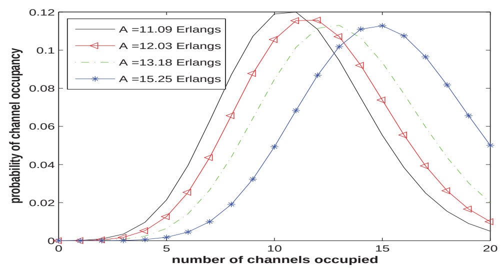


Figure C.5: Effects of traffic load, A, on the probability of channel occupancy.

As Regions W and Z are separately modeled as BD processes (described in the previous section), we can obtain the corresponding steady-state probabilities of channel occupancy. Let p_n and p_m be the probabilities that n and m channels are respectively occupied in Regions W and Z. Using Erlang's first formula [17], we have

$$p_n = \frac{A_W^n/n!}{\sum_{k=0}^{2r} A_W^k/k!}, \quad n = 0, 1, 2, \ldots, 2r, \tag{4}$$

$$p_m = \frac{A_Z^m/m!}{\sum_{k=0}^{6r} A_Z^k/k!}, \quad m = 0, 1, 2, \ldots, 6r, \tag{5}$$

where $A_W = \lambda_W/\mu_W$ and $A_Z = \lambda_Z/\mu_Z$. Here, λ_W/μ_W and λ_Z/μ_Z are the respective average traffic injected by each user in the corresponding region.

As the probabilities in Eqs. (4) and (5) are obtained from two independent regions (i.e., W and Z), the total probability, p_R, that n channels are occupied in Region-W and m channels are occupied in Region-Z becomes $p_R = p_n \cdot p_m$.

The corresponding blocking probabilities for the regions can be obtained according to the Erlang-B formula [17], as already described in Sec. IV.A.

For a given region, Fig. C.5 illustrates the behavior of the probability of channel occupancy for different values of A. As illustrated in the figure, the probability of higher number of channel occupancy is a little skewed towards right for increasing levels of A. This gives us an indication that for a given number of total available channels, with higher A, there is a higher tendency of channel occupancy in the upper half of the total channel range. Thus, there is a relatively higher chance of the system moving to a higher channel occupancy level from the lower levels. Furthermore, the channel occupancy probability drops, however less sharply, with increasing value of A near the maximum channel occupancy level. This means, e.g., for $A = 15.25$ Erlangs in the figure, the highest probability of channel occupancy is for 15 channels, and not for 10 channels (which is the middle value of channel occupancy) or for 20 channels (which is the maximum value of channel occupancy). Thus, once the transmission power of Region-W is increased (based on the channel occupancy information) to cover $cell_C$ (an action in the action space; to be explained in the next section), the probability of reducing the transmission power (another action in the action space) should be relatively low. Such information can be useful for the policy development, e.g., for assigning the probabilities of the system going from one channel occupancy interval to another in the state-transition matrix **P**.

*B. State-transition Matrix **P***

In such type of optimization problems, the state-transition matrix **P** may be obtained after gathering historical data based on the inspection results [18]. In this paper, we obtain **P** by sampling the arrival and departure statistics evolved through time for a given network. Thus, we can assign reasonably appropriate values to the elements of **P** by observing the trend of channel occupancy probabilities for different number of occupied channels.

Each state of the MC for **P** represents a bounded range for the total number of channels occupied collectively in Regions W and Z. We define two low and two high values of threshold channel occupancy in these regions respectively. Let t_{Wl_1} and t_{Wl_2} denote these lower threshold values and t_{Wh_1} and t_{Wh_2} denote the higher threshold values in Region-W, and similarly t_{Zl_1}, t_{Zl_2}, t_{Zh_1}, and t_{Zh_2} for Region-Z. We can granulate[2] the states ($\mathbf{s} = \{s_0, s_1, ..., s_5\}$) used in **P** as follows, where *occupied channels* means the total number of channels occupied collectively in Regions W and Z:

- s_0: 0 occupied channels,
- s_1: $0 <$ occupied channels $\leq t_{Wl_1} + t_{Zl_1}$,
- s_2: $t_{Wl_1} + t_{Zl_1} <$ occupied channels $\leq t_{Wl_2} + t_{Zl_2}$,
- s_3: $t_{Wl_2} + t_{Zl_2} <$ occupied channels $\leq t_{Wh_1} + t_{Zh_1}$,
- s_4: $t_{Wh_1} + t_{Zh_1} <$ occupied channels $\leq t_{Wh_2} + t_{Zh_2}$,
- s_5: $t_{Wh_2} + t_{Zh_2} <$ occupied channels $\leq 2r$.

Each element of **P** represents a probability of transition between two states. For example, an element $p_{s_0 s_1}$ of **P** means a probability of transition from s_0 to s_1, i.e., the probability that fewer than or equal to $t_{Wl_1} + t_{Zl_1}$ number of channels are occupied in the next step given that currently no channels are occupied.

[2]Here, we have used six states, i.e., $s_0, ..., s_5$. Finer granularity would lead to more accuracy but at a cost of more computational complexity.

Now, for illustration clarity, we can write **P** in a two-dimensional elemental form as p_{ij}, $i, j = 0, 1, ..., 5$. Here, i and j represent the row and column indices and respectively indicate the present and next state. Moreover, $\sum_j p_{ij} = 1$. The values to the elements of **P** can be assigned using the information obtained from Eqs. (4) and (5), and Fig. C.5.

In order for our model to be effective, there are a few necessary conditions to be met by the above mentioned thresholds, described as follows:

1) $0 < t_{Wl_1} + t_{Zl_1} < t_{Wl_2} + t_{Zl_2} \leq r$. This guarantees that the lower thresholds remain below 50% of the total allowed channel occupancy in Region-W, ensuring a fairer distribution of channels among the above mentioned six possible states of the MC.

2) $r < t_{Wh_1} + t_{Zh_1} < t_{Wh_2} + t_{Zh_2} \leq 2r$. This guarantees that the higher thresholds in Regions W and Z remain within the supportable channel limits of Region-W. Hence, Region-W is given priority over Region-Z because it is Region-W that needs to support the users of Region-Z if BS_C is switched off.

Furthermore, the steady-state values of the elements of **P**, i.e., π_i, $i = 0, 1, ..., 5$, can be easily obtained from **P** [16]. These are the limiting values of the probability of the system being in each of the states, irrespective of the initial state. For example, if p_{12} represents the transition probability to state s_2 from s_1 then π_2 is the long-term probability of the system being in state s_2. The values of π_i are used to obtain the steady-state values of the unconditional probability y_{ik} (see Eqs. (7) and (8)) from the conditional probability D_{ik} (see Eq. (6)) of action i in state k. Hence, y_{ik} are our decision variables to be optimized in the cost function for the policy (explained in Subsection V.D).

Having developed **P** and associated necessary conditions, we now move on to develop the policy to be incorporated into the FMDP. Later, we will furnish a linear program for the optimization purpose.

C. FMDP and the Policy

An FMDP is a reinforcement learning technique that satisfies the Markov property [18]. It is defined by its state and action spaces[3] and by one-step dynamic of the environment. The state space is composed of all the states that the process can assume (i.e., for example, $s_0, s_1, ..., s_5$) while the action space consists of all the possible actions that can be taken from these states. Given any state and action, there is a probability of transition to another state. Thus, a state-transition in our FMDP means that the process moves with certain probability to another state based

[3]From here on, the terms *action* and *decision* would be interchangeably used in this paper.

on the current state-action pair. Such probabilities are conditional and are generally named as transition probabilities. Furthermore, given any current state and action, together with any next state, there is an expected value of the associated cost (due to the action taken). This cost, e.g., in our problem, is the transmission power consumption (in Watts) when another action is taken based on the current state and action. This action may be to switch to the PS-ON or the PS-OFF mode, thereby, increasing or decreasing the transmission power consumption. As the actions are taken with certain probabilities, our aim in this scheme is to minimize this *expected* cost. Therefore, we develop a policy for this purpose.

The policy is meant to determine a probability distribution for the actions k ($k = 1, ..., K$), to be taken when the system is in state i ($i = 0, 1, ..., M$). Thus, there are a total number of K possible actions and $M + 1$ possible states. Hence, we define the conditional probability distribution as

$$D_{ik} = Prob\{action = k \mid state = i\}, \tag{6}$$

where $\sum_k D_{ik} = 1$ and $0 \leq D_{ik} \leq 1$.

The action space[4] is defined as follows:

- Action 1: Go to the PS-ON mode, i.e., shut down BS_C and increase transmission power of Region-W,
- Action 2: Go to the PS-OFF mode, i.e., switch on BS_C and decrease transmission power of Region-W,
- Action 3: Keep the current power mode unchanged.

It is worth mentioning that each of these actions can be taken with certain probability based on the current state and previous action. An action may or may not lead the system to a new state. If the system goes to a new state, the process may or may not go to another state based on the new action-state pair, and so on.

Hence, our policy can be characterized by a policy matrix, **D**, where each state of **D** is one of the states defined for **P** and each action is one of the actions from the above mentioned action space. Therefore, each element D_{ik} of **D** represents the probability of action k in state i; and each row of **D** is a Probability Distribution Function (PDF) of all the possible actions that can be taken in a state. For example $i = 0$ means state s_0 and $k = 1$ means Action 1; and so on.

[4]For the sake of simplicity, the action space has been deliberately kept small. Larger action space leads to more complex policy, in general.

A simple example of a policy matrix, **D**, is shown below, given that there are 6 states and 3 actions.

$$\mathbf{D}= \begin{array}{c} \\ s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{pmatrix} p_h & 0 & 1-p_h \\ p_h & 0 & 1-p_h \\ p_m & 0 & 1-p_m \\ p_l & p_l & 1-2p_l \\ 0 & 1-p_h & p_h \\ 0 & p_h & 1-p_h \end{pmatrix} \end{array},$$

where the subscripts, h, m, and l of p, respectively indicate the high, medium, and low probability of an action in a state. Actions 1, 2, and 3 are respectively represented by a_1, a_2, and a_3. As an example, the top-left element of **D** means that given no channels are occupied, there is a high probability that the system goes to the PS-ON mode.

D. Cost Optimization and Linear Programming

In order to optimize the transmission power consumption cost through a linear program, we need first to define the decision variables, y_{ik}, to be optimized. Let y_{ik} be the steady-state probability that the system is in state i and action k is taken, i.e.,

$$y_{ik} = Prob\{action = k \text{ } and \text{ } state = i\}. \tag{7}$$

From the rules of conditional probability, y_{ik} and D_{ik} can be related[5] as

$$y_{ik} = \pi_i D_{ik}. \tag{8}$$

The long-term expected transmission power consumption cost is given as

$$E(C) = \sum_{i=0}^{M} \sum_{k=1}^{K} \pi_i C_{ik} D_{ik} = \sum_{i=0}^{M} \sum_{k=1}^{K} C_{ik} y_{ik}, \tag{9}$$

where C_{ik} denotes the cost (in Watts) incurred when action k is taken in state i, and it can be calculated using Eq. (12).

As mentioned earlier, our goal is to minimize the cost function, $X \equiv E(C)$. The minimization of X means finding optimal values of the decision variables y_{ik} which minimize the expected cost (see Eqs. (7) and (9)). In order to do so, the model is

[5] Note the difference between Eqs. (6) and (7). The former is a conditional probability while the latter is a limiting unconditional probability.

turned into a linear program as follows:

$$\begin{aligned} \min \quad & X = \sum_{i=0}^{M} \sum_{k=1}^{K} C_{ik} y_{ik} \qquad (10) \\ \text{s.t.} \quad & \sum_{i=0}^{M} \sum_{k=1}^{K} y_{ik} = 1, \\ & \sum_{k=1}^{K} y_{jk} - \sum_{i=0}^{M} \sum_{k=1}^{K} y_{ik} p_{ij}(k) = 0, \; j = 0, 1, ..., M, \\ & y_{ik} \geq 0 \; for \; i = 0, 1, ..., M \; and \; k = 1, 2, ..., K. \end{aligned}$$

In this linear program, argument k in $p_{ij}(k)$ indicates that the appropriate transition probability depends on action k. The first constraint is necessary because as y_{ik} are probabilities of all action-state pairs, their sum must be 1. Hence, this is the normalization equation for the Markov process. The second constraint indicates that the steady-state probability of being in state j is the same as the probability computed by conditioning on the state and action chosen one stage earlier (as per definition of an MC). Hence, this is the balance equation for the Markov process. The third constraint imposes a lower bound on the decision variable values (as they represent probabilities, they should be non-negative).

VI. BASE STATION POWER CONSUMPTION

The power consumption of a BS mainly consists of two major components, i.e., a component responsible for the total transmission power of the emitted signal and a component for the fixed power consumption. Below, we first describe how to obtain the total transmission power for a BS. The fixed power consumption component is later accounted for.

Quality of service can be taken as the acceptable cumulative effect on user satisfaction of all imperfections affecting the service. To ensure a certain QoS level for any specific type of service, a BS is required to maintain a minimum level of Signal to Interference and Noise Ratio (SINR) for a target MS. Hence, for a given noise level, there should be a minimum level of received power, P_{rx}, at the MS. Therefore, in order to ensure this minimum P_{rx}, the BS needs to transmit at least with a power level of P_{tx}. Thus we can write

$$P_{tx} = P_{rx} \cdot d^{\alpha} \cdot L_1. \qquad (11)$$

Here $P_{rx} = \gamma \cdot (W \cdot T_0 + I)$, where γ is the required minimum SNR level for the target MS, W is the channel bandwidth, T_o is the thermal noise level, and I is the maximum assumed value of the interference level. Furthermore in Eq. (11), the distance between the MS and the BS is denoted by d, and α is the path-loss coefficient. The losses due to fading and building penetration etc are collectively given by component L_1. The total transmission power consumption by the BS towards a total number of N_A active users in a cell/sector can thus be written as

$$P_{Tot\_tx} = N_A \cdot P_{tx}. \tag{12}$$

Eq. (12) is calculated for the worst case power consumption scenario, i.e., the MSs are considered to be located at the cell/sector boundaries.

The fixed power consumption component, P_{fix}, of a BS includes power consumed due to heating effects, electronic processing of transmitted and received signals, etc. For the sake of simplicity, it is assumed as constant.

Therefore, the total BS power consumption can now be written as

$$P_{BS} = L_2 \cdot N_S \cdot P_{Tot\_tx} + P_{fix}, \tag{13}$$

where scalar N_S represents the number of sectors in a cell. The losses associated with components of a BS, e.g., antenna feeder cable loss, directional antenna gain, power amplifier efficiency etc are collectively denoted as L_2 in Eq. (13).

In the deterministic scheme case, the number of users can be obtained based on the channel threshold values. Similarly, the number of users for the FMDP-based scheme can be obtained based on the state information the system is currently operating in. Then using the value of N_A (i.e., the number of active users), Eq. (12), and Eq. (13), we can obtain the total transmission power consumption of BS for a sector (and for a cell when scaled accordingly).

VII. Numerical Results: Sojourn Time-based Deterministic Scheme

In this and the next sections, we evaluate numerically the performance of the proposed schemes using MATLAB. The evaluation parameters are summarized in Table C.1 [2], [19], and [20].

The total transmission power of a cell before and after the increase in transmission power of the sector antenna is calculated. Similarly the total consumed power of the network before and after the switch-off of BS_C calculated. Thus, the benefit as well as the tradeoff in the consumed power are obtained.

As mentioned earlier, the network power has two components, i.e., the power consumed for transmission of signals, and the fixed BS power consumption component. In the following, we discuss the effects of our scheme on both of these

Table C.1: Analysis parameters.

Parameter	Value
Channel bandwidth W	5 MHz
Minimum SNR value for a service γ	-18 dB
Total noise + interference density	-166 dBm/Hz
Receiver sensitivity	-117 dBm
Loss component (Rayleigh fading)	$2 \sim 5$ dB
Loss component (building penetration)	$12 \sim 15$ dB
Loss component (shadowing)	$6 \sim 7$ dB
Path-loss coefficient α	4
Distance between MS and BS d	$(450, 900)$ meters
BS directional antenna gain	10 dB
MS antenna gain	0 dB
P_{fix}	60 W
BS antenna feeder cable loss	-2 dB
Power amplifier efficiency	50%
Number of channels in each sector r	15
Average injected traffic by each user λ/μ	$(0.4, 0.5, 0.6)$ Erlangs
GoS level	$(0.5, 1, 2)$ %

components for different values of GoS, $t_l + t_l'$, $t_u + t_u'$, and λ/μ. In both Figs. C.6 and C.7, the upper part illustrates total transmission power and the lower part illustrates total network power consumption for the sojourn time-based scheme.

The bars in Set 3 of the upper parts of Figs. C.6 and C.7 indicate the extra cost in terms of transmission power in the PS-ON mode versus the PS-OFF mode. Similarly, the bars in Set 3 of the lower parts of all of these figures indicate the saving in the network power.

It is worth mentioning that all these results are for the transmission and total power consumption of the network. In other words, MS power consumption is not included in these figures. GoS requirement is given as a percentage value and indicates the blocking probability.

A. Network Transmission Power Consumption

Here, we present the effects of GoS and λ/μ; hysteresis region boundary shifting; and hysteresis region length variation on the network transmission power consumption. Moreover, the effects on the total network power consumption are also described.

A.1. Effects of GoS and λ/μ

As illustrated in Fig. C.6, for λ/μ =0.4, loose requirement on GoS (e.g., 2%) causes more transmission power consumption in both the PS-ON and the PS-OFF modes. This is because that the network can provide service to a higher number of

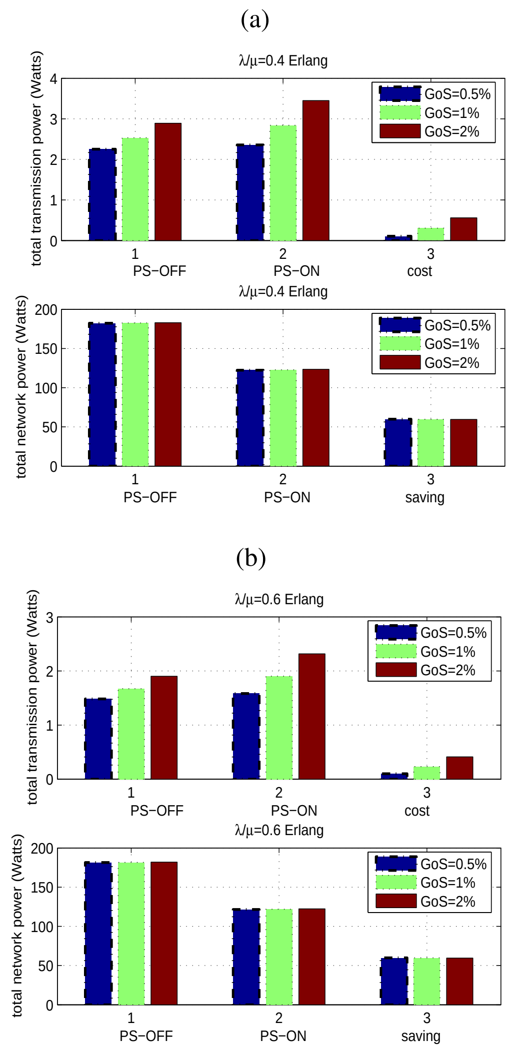


Figure C.6: Deterministic scheme: Total transmission and network power consumptions for (a) $\lambda/\mu = 0.4$, $t_l + t_l' = 5$, $t_u + t_u' = 10$; (b) $\lambda/\mu = 0.6$, $t_l + t_l' = 5$, $t_u + t_u' = 10$.

users if GoS requirement is relaxed. That is, fewer number of calls are blocked with higher GoS value. Furthermore, the PS-ON mode consumes more transmission power than the PS-OFF mode. This is because of the reason that in the PS-ON mode, two of the sector antennas (i.e., of BS_A and BS_B) need to cover $cell_C$ as well. Hence, they are required to transmit with higher power to reach the users at distance $2d$, causing further power consumption due to the exponential relation between transmission power and distance. This indicates the extra cost in terms of transmission power that the network operator has to bear in order to save network power.

Comparing Figs. C.6(a and b) (upper parts only), similar trend for transmission power consumption is observed. However, the magnitude of transmission power consumed for any GoS requirement in Fig. C.6(b) is lower than the corresponding values in Fig. C.6(a). For example, in the PS-ON mode, for the same value of GoS (here, GoS = 2%), the total transmission power is about 2.4 Watts for $\lambda/\mu = 0.6$ in Fig. C.6(b); whereas it is more than 3 Watts for $\lambda/\mu = 0.4$ in Fig. C.6(a). This is because of the reason that as λ/μ is taken 0.6 in Fig. C.6(b), the amount of traffic injected by each user into the network is higher. This leads to a lower number of supported users at a given time because a higher number of channels is occupied by each user on average. A similar trend can also be observed for the PS-OFF mode.

A.2. Effects of Hysteresis Region Boundary Shifting

By comparing Figs. C.7(a and b), we observe a similar trend as discussed in the above paragraphs for Fig. C.6. However, a higher amount of transmission power is consumed in the PS-ON mode for all mentioned GoS requirements than the corresponding values in Figs. C.6, respectively. This is because that since, in Figs. C.7(a and b), the thresholds $t_l + t_l'$ and $t_u + t_u'$ are increased by one, there is a higher number of states in the pure PS-ON region (as illustrated in Fig. C.4 and Sec. IV.B). This implies a higher amount of total expected time spent in the PS-ON mode. Hence, the network remains in the PS-ON mode for a relatively longer fraction of time than in the PS-OFF mode. Furthermore, in the PS-ON mode, since two sector antennas (i.e., of BS_A and BS_B) transmit at longer distance (i.e., $2d$), there is higher transmission power consumption even though the antennas of BS_C are not transmitting. Thus, for a given λ/μ and GoS requirement, *shifting the thresholds to the higher values results in a higher transmission power consumption in the PS-ON mode.*

Furthermore, comparing Figs. C.7(a and b) with Figs. C.6(a and b) respectively, for the PS-OFF mode, little difference in transmission power consumption for the same λ/μ is observed. The reason for this effect is because that, as in the PS-OFF mode all the antennas are transmitting at normal distance, shifting the hysteresis region to just one higher value (and thus reducing the pure PS-OFF region) only decreases the transmission power slightly.

B. Total Network Power Consumption

The effects of the proposed scheme on the total network power consumption are illustrated in the lower parts of Figs. C.6 and C.7. As can be observed in all the figures, *the proposed scheme saves a considerable amount of network power for all sets of GoS, λ/μ, $t_l + t_l'$, and $t_u + t_u'$.* In all cases, the amount of power saved is more than 50 Watts out of about 185 Watts. It is worth mentioning that the values

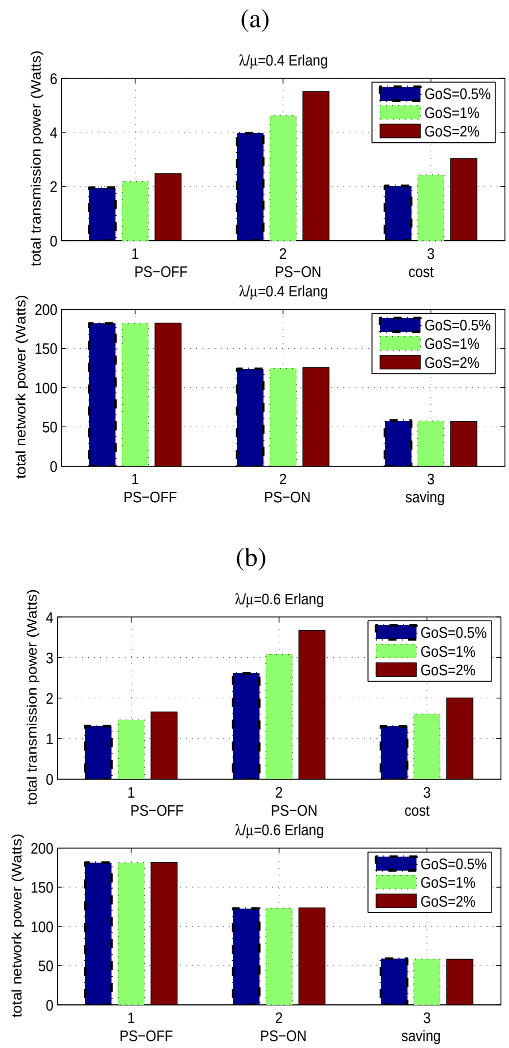


Figure C.7: Deterministic scheme: Total transmission and network power consumptions for (a) $\lambda/\mu = 0.4$, $t_l + t_l' = 6$, $t_u + t_u' = 11$; (b) $\lambda/\mu = 0.6$, $t_l + t_l' = 6$, $t_u + t_u' = 11$.

of saved network power for different parameters are somewhat different. However, this difference is not clearly visible in the figures due to the scale on vertical axes.

The major reason for the large amount of network power saving is due to BS switch-off. Even when the transmission range of an antenna (i.e., for example, in Region-W) is increased from d to $2d$ while the remaining sector antennas of the corresponding BS are transmitting at distance d, the total transmission power value remains much lower than the fixed power consumption component of a BS, due to lower traffic intensity in Region-C. Hence, a BS's fixed power consumption always overwhelms the corresponding transmission power consumption for reasonable values of d. Furthermore, when a BS is switched off (i.e., BS_C in this case), its own transmission power is also saved. Thus, switching off just one BS in our network results in a large amount of network power saving. Nevertheless, this saving is achieved at a cost of transmitting extra power from two sector antennas (of BS_A and BS_B). Numerically with our scheme, *more than 50 Watts of network power can be saved by sacrificing about 4 - 5 Watts in terms of extra transmission power.* This implies a cost-to-benefit ratio of about 1:10.

VIII. NUMERICAL RESULTS: FMDP-BASED PROBABILISTIC SCHEME

This section presents the results achieved by numerical evaluation of the FMDP-based probabilistic scheme. The linear program is solved by using the well-known simplex algorithm[21].

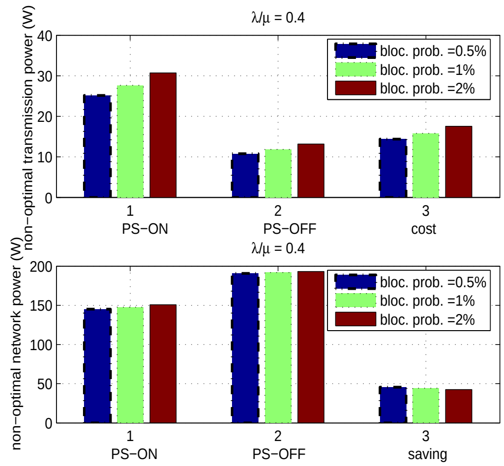
A. Power Consumption without Optimization

For a fair comparison with the obtained optimization results, we first need to know the non-optimal power consumption picture for our FMDP-based scheme. With non-optimal, we mean the normal power consumption without minimization algorithm applied on the system, neither in the PS-ON nor in the PS-OFF mode. The upper and lower parts of Fig. C.8(a) present, respectively, the non-optimal network transmission and total power consumption in both the PS-ON and the PS-OFF modes for $\lambda/\mu = 0.4$ and different levels of blocking probability. Similarly, Fig. C.8(b) illustrates the same set of results for $\lambda/\mu = 0.6$.

As illustrated in Fig. C.8(a) (the upper part), the values of power consumption are lower in the PS-OFF mode than the corresponding values in the PS-ON mode. This result is expected and it is because that in the PS-ON mode, two of the sector antennas (i.e., of $cell_A$ and $cell_B$) are required to transmit with a higher power level to reach the users in $cell_C$ at distance $2d$. This causes further power consumption due to the exponential relation between the transmission power and the distance (as implied in Eq. (11)).

Moreover, by comparing the upper and lower parts in Fig. C.8(a), we observe a clear advantage in terms of total network power consumption. That is, *the system can save a considerable amount of network power in the PS-ON mode at certain extra cost of the transmission power*. This is because that in the PS-ON mode, only two BSs are consuming power and the BS_C is switched off. This saves a reasonable amount of network power because P_{fix} for BS_C as well as its antenna transmission power are not contributing to the network power consumption. However, comparing Set 3 in the upper and lower parts of Fig. C.8(a), we observe that the ratio of the network power saving to the extra transmission power cost decreases. This is because that when the transmission power is higher, there is a relatively lower amount of saved total network power.

(a)



(b)

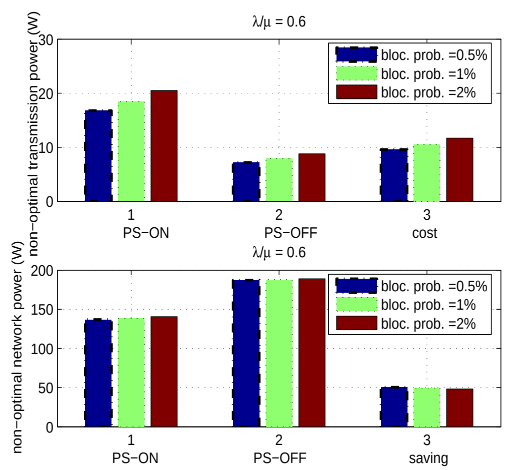


Figure C.8: Probabilistic scheme: Non-optimal network transmission and total power consumption for (a) $\lambda/\mu = 0.4$, and (b) $\lambda/\mu = 0.6$.

Fig. C.8(b) illustrates the total non-optimal transmission and network power consumption given $\lambda/\mu = 0.6$. The trend of power consumption is in accordance with that of Fig. C.8(a). However, the values of the consumed network power (both transmission and total) are lower in Fig. C.8(b) compared with those in Fig. C.8(a). This effect is due to the reason that λ/μ is higher in Fig. C.8(b), meaning that more traffic is injected by each user on average. Hence, a fewer number of users can be supported for a given number of occupied channels, resulting in a lower amount of transmission (and network) power consumption.

B. Optimum Power Consumption

The optimized[6] transmission power consumption of the system for different

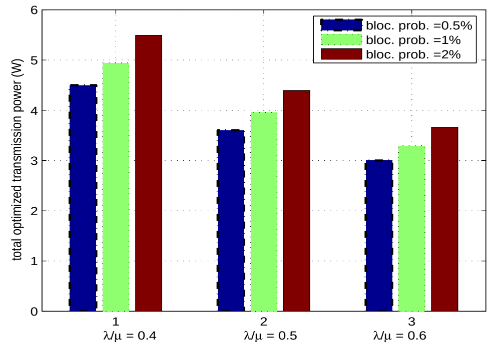


Figure C.9: Optimum network transmission power consumption.

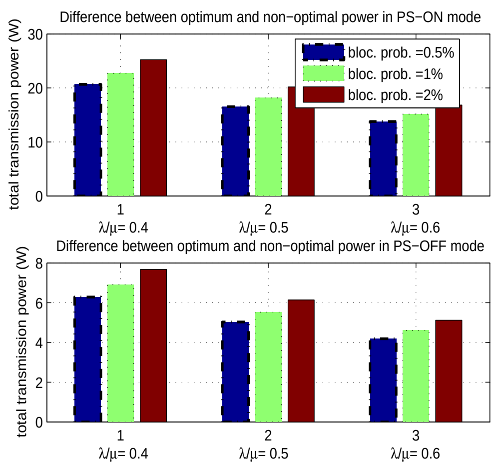


Figure C.10: Difference between non-optimal and optimum transmission power consumption in the PS-ON mode (upper part) and the PS-OFF mode (lower part).

values of GoS level and λ/μ is illustrated in Fig. C.9. The linear program (Eq. (10) and associated constraints) and Eq. (9) are respectively used to obtain the minimized y_{ik} and the minimized expected transmission power consumption cost. As expected, the optimized (i.e., minimized) transmission power values for lower values of λ/μ are higher for given blocking probability levels. This is because that for higher λ/μ, on average, more traffic is injected by individual users. Thus a higher number of channels is more quickly occupied, pushing the system towards the PS-OFF mode. Hence, a lower amount of transmission power is consumed because antennas only need to transmit over distance d.

A comparison of Fig. C.9 with the upper parts of Fig. C.8(a and b) illustrates that the optimized transmission power consumption is the lowest. This is because

[6] As the goal of the optimization is to minimize only the transmission power consumption, the network power consumption results will not be discussed.

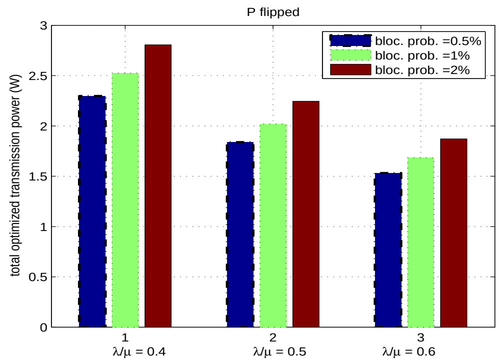


Figure C.11: **P**-flipped: Optimum network transmission power consumption.

that *instead of switching to the PS-ON/OFF mode definitely, we switch to the PS-ON/OFF mode with certain probability* depending on the state-action pair in the FMDP, so the power consumption is optimized. This ensures minimum long-term transmission power consumption due to the time intervals when very few channels are occupied.

In Fig. C.10 (the upper part), the difference between the non-optimal transmission power consumption in the PS-ON mode and the optimum transmission power is illustrated. That is, e.g., Set 3 in Fig. C.10 (upper part) illustrates the corresponding difference between Set 1 of Fig. C.8 (upper part) and Set 3 of Fig. C.9. The trend for this result is in accordance with the results described in the above paragraphs of this section. Similarly, Fig. C.10 (lower part) illustrates the same for the PS-OFF mode. That is, e.g., Set 3 in Fig. C.10 (lower part) illustrates the corresponding difference between Set 2 of Fig. C.8 (upper part) and Set 3 of Fig. C.9. From these results, we observe that by using the proposed optimization scheme, even in the PS-OFF mode, a further saving of about 5 Watts is achieved in terms of total transmission power, for the given levels of blocking probability and λ/μ.

B.1 Effects of Flipped ***P***

Figs. C.11 and C.12 respectively present the same phenomena as those in Fig. C.9 and C.10, however, for a flipped version of **P** along its rows. Flipping **P** along its rows has the effect of reversing the channel occupancy probabilities from states s_0 to s_5. That is, the steady-state probability of s_5 in **P** becomes steady-state probability of s_0 in the flipped **P**, and so on. As expected, the trends are similar if we compare Fig. C.9 with Fig. C.11 or Fig. C.10 with Fig. C.12. However, as observed by comparing Fig. C.11 with Fig. C.9, the optimized transmission power consumption is lower in Fig. C.11. The reason for this effect is because that as **P** is flipped, the transition probabilities between states are also changed. For example, if the transition probability between s_0 and s_4 was high in the first place, it becomes low

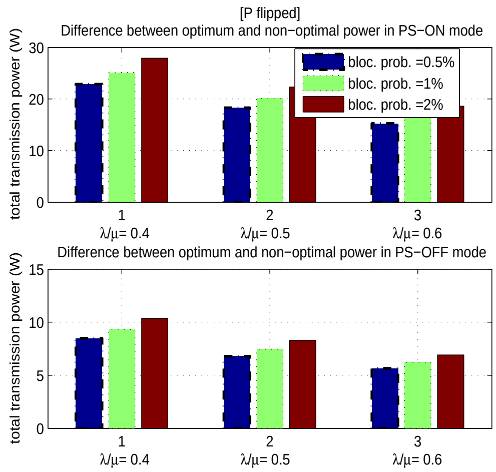


Figure C.12: **P**-flipped: Difference between non-optimal and optimum transmission power consumption in the PS-ON mode (upper part) and the PS-OFF mode (lower part).

in the flipped **P**. Therefore, arguing on these lines, in general in the flipped **P**, the probabilities that a higher number of channels are occupied are relatively lower than probabilities that a lower number of channels are occupied. Thus, a flipped version of **P** results in an even lower amount of optimized transmission power.

However, it is worth mentioning that this flipped version of **P** is discussed here merely for the purpose of comparison. Furthermore, this also indicates that the optimization policy may be improved further based on whether the system operator requires higher/lower probabilities for the states in **P**. How such policies can be formed and improved is, nevertheless, beyond the scope of this paper.

IX. Comparison and Future Discussions

A comparison between Figs. C.6 and C.8 indicates that, for transmission power consumption, the values in any mode for the FMDP-based scheme (for non-optimal power only) are higher than those for the sojourn time-based scheme, for given values of λ/μ. The reason for this effect is due to the fact that in the sojourn time-based scheme, the power consumption is associated with the values of the thresholds $t_l + t_l'$ and $t_u + t_u'$. Since the values of $t_l + t_l'$ are much lower than the maximum value of r, a few channels are occupied meaning that a few users are active. So the transmission power consumption value is also correspondingly low.

Intuitively, one may expect that the optimized transmission power would be the least when the two schemes are compared. Surprisingly, by comparing Fig. C.9 with any of the Figs. C.6 and C.7, we observe that the optimized transmission power consumption is in fact relatively higher. The reason is, again, due to the lower values of the thresholds ($t_l + t'_l$ and $t_u + t'_u$) selected in our earlier work [12] which bring the power consumption in the PS-ON and the PS-OFF modes lower, compared with the optimum values of power in the FMDP-based scheme. However, the selection of higher values of $t_l + t'_l$ and $t_u + t'_u$ will generate higher transmission power consumption in the sojourn time-based scheme than the optimum consumption generated in the FMDP-based scheme.

Furthermore, as power consumption is sensitive to the selection of optimum threshold levels as well as the hysteresis region length, they need to be explored further for an optimum level of transmission power consumption. On the other hand, in the FMDA-based scheme, if the lower channel occupancy levels are assigned with higher probabilities (in **D**), the corresponding total transmission power consumption will be lower due to the fact that the system spends more time in low channel occupancy states. Therefore, a joint optimization, targeted at threshold selection in the deterministic scheme as well as probability assignment to channel occupancy states in the probabilistic scheme, is required.

Moreover, as our proposed schemes are not restricted to any particular time intervals of a day, they are applicable to all times and all traffic intensities. That is, during the rush hours the schemes may more often go to the PS-OFF mode, and during the low-traffic hours they more often remain in the PS-ON mode. Thus they can save power according to the traffic intensities at all times than just low-traffic times of a day.

To summarize, we claim that both proposed schemes are useful for network operators. The sojourn time-based is preferable in reducing the network transmission power consumption. The FMDP-based scheme can be preferable by the operators in situations where estimating the long-term power consumption cost of their networks is a priority.

X. Conclusions and Future Work

In this paper, we have proposed two independent teletraffic-based schemes to analyze the transmission and network power consumption for a simple network with micro-cells. Numerical results demonstrate that in the sojourn time-based scheme, our hysteresis-based approach saves overall network power for an acceptable amount of extra transmission power consumption cost. In the FMDP-based scheme, the analysis can be utilized to minimize the long-term transmission power consumption in a cellular network.

As our future work, we will extend the current scenario to a more realistic network with multi-tier cells. As in such a scenario, more than one cells can enter into sleep mode, we will investigate the optimum number and locations of the switched-off cells and compensating sectors. Hence, interference from other cells, inter-cell interaction, co-frequency deployment etc will also be thoroughly investigated.

Another further extension of this work will be the inclusion of IP-traffic based analysis with asymmetric traffic distribution and non-uniform MS distribution.

REFERENCES

[1] J.T. Louhi, "Energy Efficiency of Modern Cellular Base Stations," *in Proc. International Telecommunications Energy Conference (INTELEC)*, Rome, Italy, October 2007.

[2] iCell COMPAC IP-RAN EV-DO NETWORK: Technical Specifications. Homepage: http://www.starsolutions.com/index.php/radio-access-network-components/bts-compact-micro-s

[3] *Node B datasheets*, http://www.motorola.com/, 2008

[4] "Vodafone Using Ericssons New Power-saving Base Station Feature," http://www.3g.co.uk/PR/Dec2007/5524.htm, (website accessed on Sept. 02, 2011).

[5] B. Debaillie, A. Giry, M. J. Gonzales, L. Dussopt, M. Li, D. Ferling, and V. Giannini, "Opportunities for Energy Savings in Pico/Femtocell Base-Stations, " *in Proc. Future Network and Mobile Summit*, Warsaw, Poland, June 2011.

[6] S. Bhaumik, G. Narlikar, S. Chattopadhyay, and S. Kanugovi, "Breath to Stay Cool: Adjusting Cell Sizes to Reduce Energy Consumption," *in Proc. First ACM SIGCOMM Workshop on Green Networking*, New Delhi, India, August 2010.

[7] M. A. Marsan and M. Meo, "Energy Efficient Management of Two Cellular Access Networks," *in Proc. GreenMetrics 2009 Workshop*, Seattle, WA, USA, June 2009.

[8] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal Energy Savings in Cellular Access Networks," *in Proc. GreenComm'09 - First International Workshop on Green Communications*, Dresden, Germany, June 2009.

[9] M. A. Marsan, L. Chiaraviglio, and D. Ciullo, "A Simple Analytical Model for the Energy-efficient Activation of Access Points in Dense WLANs," *in Proc.*

1st International Conference on Energy-efficient Computing and Networking, e-Energy 2010, New York, USA, April 2010.

[10] J. Zhou, M. Li, L. Liu, X. Sge, and L. Chen, "Energy Source Aware Target Cell Selection and Coverage Optimization for Power Saving in Cellular Networks," *in Proc. IEEE/ACM GreenCom-CPSCom*, Hangzhou, China, December 2010.

[11] Z. Chong and E. Jorswieck, "Analytical Foundation for Energy Efficiency Optimisation in Cellular Networks with Elastic Traffic," *in Proc. of 3rd International ICST Conference on Mobile Lightweight Wireless Systems (Mobilight)*, Bilbao, Spain, May 2010.

[12] Z. H. Abbas and F. Y. Li, "A Novel Teletraffic-based Power Saving Scheme for Cellular Networks with Microcells," *in Proc. 7th IEEE Performance and Management of Wireless and Mobile Networks P2MNET Workshop in conjunction with 36th IEEE LCN*, Bonn, Germany, October 2011.

[13] W. C. Y. Lee, "Overview of Cellular CDMA," *IEEE Trans. on Vehicular Technology*, Vol. 40, No. 2, May 1991.

[14] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd Edition, Prentice-Hall Inc, 2002.

[15] W. J. Stewart, *Probability, Markov Chains, Queues, and Simulation: the Mathematical Basis for Performance Modeling*, Princeton University Press, 2009.

[16] S. M. Ross, *Introduction to Probability Models*, 9th edition, Elsevier Inc., 2007.

[17] A. Popoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th Edition, McGraw-Hill Companies, Inc., New York, USA, 2002.

[18] F. S. Hillier and G. J. Lieberman, *Introduction to Operations Research*, 8th edition, McGraw-Hill Companies Inc., New York, USA, 2005.

[19] T. Ojanperä and R. Prasad, *Wideband CDMA for Third Generation Mobile Communications*, Artech House, 1998.

[20] B. T. Ahmed, "WCDMA downlink capacity of cigar-shaped microcells for underground metro service," *European Trans. on Telecommunications,*, Vol. 20, No. 4, June 2009.

[21] I. Maros, *Computational Techniques of the Simplex Method*, Kluwer Academic Publishers, The Netherlands, 2003.

Paper D

Analysis of Mobile Station Battery Energy Consumption in Heterogeneous Networks using Markov Processes with State Transition Rewards

---

Title:	Analysis of Mobile Station Battery Energy Consumption in Heterogeneous Networks using Markov Processes with State Transition Rewards
Authors:	Ziaul Haq Abbas and Frank Y. Li
Affiliation:	Department of Information and Communication Technology, University of Agder (UiA), N-4898 Grimstad, Norway
Conference:	To be submitted to *IEEE International Conference on Communications (ICC 2013)*, Budapest, Hungary, June 2013.

Analysis of Mobile Station Battery Energy Consumption in Heterogeneous Networks using Markov Processes with State Transition Rewards

Ziaul Haq Abbas and Frank Y. Li
Department of Information and Communication Technology
University of Agder (UiA), N-4898 Grimstad, Norway
E-mails: {ziaul.h.abbas, frank.li}@uia.no

Abstract — **With the increasing ubiquity of heterogeneous networks and the storming popularity of content-rich applications into the mobile market, reducing mobile station battery energy consumption has become an urgent requirement for both operators and manufacturers. In this paper, we propose a discrete-time Markov chain based approach to estimate the long-term energy consumption of a mobile station operating in a heterogeneous network comprising of WLAN, ad hoc, and cellular networks. Moreover, a finite Markov decision process is utilized to minimize the overall long-term average energy consumption of the battery through a policy improvement methodology. Numerical results demonstrate that by using our proposed approach, the battery energy consumption can be reduced due to the feasibility of selecting the least energy consuming alternative during the progressive policy improvement procedure. With this approach, we can also predict the long-term battery energy consumption of a mobile station.**

I. Introduction

Today's cellular networks offer not only voice services but also data and multimedia services through Internet access. At the same time, with the widespread deployment of WiFi hotspots in places like offices, airports, train stations etc, the availability of Internet access has become ubiquitous. Due to these reasons, most of the mobile devices nowadays are equipped with more than one wireless interfaces, e.g., WiFi, 3G, and GSM; and are operated in heterogeneous environments. However, as Mobile Stations (MSs) are generally battery powered, they have limited energy storage. Therefore, estimating and reducing battery energy consumption of MSs while supporting heterogeneous connections have become of great importance.

In the recent literature, there are quite a few proposed approaches which investigate energy consumption in MSs operating in hybrid/heterogeneous networks.

For example, in [1], a measurement based study is carried out for energy consumption in mobile phones for more than one type of mobile networking technologies; and consequently, a model for calculating consumed energy by network activity for each technology is developed. Using this model, the authors develop a protocol that reduces energy consumption of common mobile applications. However, their model does not apply to long-term energy usage prediction in a heterogeneous environment. In our earlier work [2], we present a distance-based power consumption analysis for MSs in a heterogeneous wireless network. The results demonstrate that for various distances between transceivers in such networks, sending and receiving data through distinct links (e.g., sending through a WLAN link and receiving through a cellular link) can help in reducing the average energy consumption of MSs, and thereby increase the corresponding battery lifetime. However, what will be the long-term energy consumption picture for MSs operating in heterogeneous fashion in such networks is not addressed there. Moreover, a novel semi-Markov analytical model of state transitions to evaluate energy consumption of MSs for instant messaging services is developed in [3]. This is an important step for MS energy consumption analysis, however, it applies only for a single application (i.e., instant messaging). Again, the impact of the developed model on the overall long-term energy consumption of MSs is not presented in their work. In [4], a general mathematical framework built on Markov Decision Process (MDP) is presented to optimize energy efficiency of mobile phones, by taking talk-time as the primary parameter. The focus of the authors includes, nevertheless, reduction of the decision table size for certain applications, for example data synchronization, in order to reduce energy consumption. However, although many approaches exist for MS energy consumption calculations in mobile and wireless networks, little previous work which analyzes MS energy consumption with heterogeneous connections using Finite Markov Decision Process (FMDP) with rewards has been reported. This observation indeed motivates our work in the paper.

In this paper, a novel link transition-based energy consumption analysis using Markov process is performed for an MS operating in an HN which is comprised of a WLAN, an ad hoc, and a cellular network. As the first step, a Discrete-Time Markov Chain (DTMC) with *rewards* is employed to analyze the asymptotic value of MS energy consumption given that it starts from a certain network link. A reward in our scenario means the value of energy consumption of the MS for operating in a given link. Afterwards, an FMDP is employed to provide the MS with more options to select the least energy consuming alternative. Therefore, the main objective of the FMDP-based analysis is to minimize the expected reward. The numerical results il-

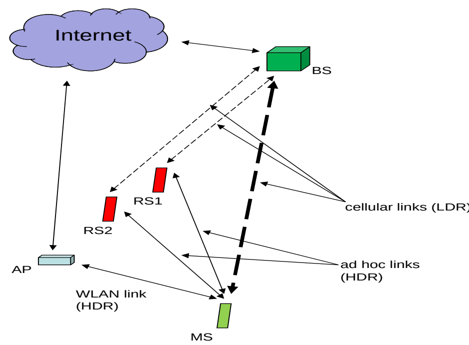


Figure D.1: Network scenario for MS with heterogeneous connections.

lustrate that by using the FMDP, the MS can progressively converge to an optimized policy to ensure that the overall long-term energy consumption is minimized. The results also reflect on how battery energy is consumed in such scenarios.

The rest of the paper is organized as follows. Sec. II describes the network scenario together with a few assumptions. Sec. III presents the analyses of the MS energy consumption with state transitions using the DTMC and FMDP processes as well as the policy iteration algorithm. Afterwards, numerical results and discussions are presented in Sec. IV, before the paper is concluded in Sec. V.

II. Network Scenario and Assumptions

The network scenario studied in this paper is illustrated in Fig. D.1. The overall Heterogeneous Network (HN) supports three types of connections, i.e., the High Data Rate (HDR) ad hoc component, the HDR WLAN component, and the Low Data Rate (LDR) cellular component. The MS can access the Internet via the Base Station (BS) either directly through the cellular link or indirectly through a one-hop ad hoc Relay Station (RS). Thirdly, the MS can also access the Internet via the Access Point (AP) through a WLAN link. Hereafter, the links of the MS through AP, BS, and RS are respectively referred to as *WLAN link*, *cellular link*, and *ad hoc link*. Through them, the MS can enjoy heterogeneous connections, but not simultaneously.

As the focus in this study is about the battery energy consumption of the MS, the AP and the BS are considered to be powered by Alternating Current (AC) power lines, and reach the Internet through wired connections. The RS and the MS are, nevertheless, considered as hand-held devices with limited battery energy.

Moreover, we assume that all involved devices support both the Point Coordination Function (PCF) and the Distributed Coordination Function (DCF) [5], for radio channel access in the WLAN link and the ad hoc link, respectively. The power consumed for switching between the links is assumed to be negligible. Furthermore, the power consumption of the MS in the processes of association/disassociation with the AP is also ignored.

For calculation convenience, we consider only one transmission cycle. This means, for example in the HDR ad hoc connection, the total duration starting from the DIFS until the ACK is received is regarded as one transmission cycle. The RS has sufficient buffer size to store the incoming data before forwarding it to the next station. Furthermore, the control and data frames will not be lost and are received as error-free packets.

The analysis is based on the assumption that only the AP can poll the MSs (because for an MS, receiving *Beacon* frames consumes much less power than periodically transmitting polling frames in an active manner). The MSs do not miss any *Beacons*, i.e., they always wake up immediately before each *Beacon* arrival. The contention-free parameter set elements [6] are included in the *Beacon* frame to keep all hand-held devices apprised of contention free operation duration. In order to use the AP for Internet access, an MS receiving this *Beacon* will set its Network Allocation Vector (NAV) to the maximum duration of Contention Free Period (CFP) to block any DCF-based access to the medium. Moreover, the Contention Period (CP) does not overrun the CFP and is long enough for the transfer of at least one maximum-size data frame and its associated control and acknowledgment frames.

For the simplicity of the analysis, ad hoc connection through only one RS is considered. A detailed analysis of the cellular as well as the PCF- and DCF-based MS energy consumption can be found in our earlier work [7].

III. Analysis of the Scenario with Markov Process

In this section we analyze the MS energy consumption in the studied scenario using Markov process. Initially, to build a mathematical foundation for our later analysis, the scenario is analyzed with a simple DTMC with rewards using a finite number of states. This gives us the MS energy consumption behavior, nevertheless, with predefined probability and fixed direction of state transitions. Later on, an FMDP is employed to analyze the same scenario but with more than one possible actions in each state of the process. This enables us to assign more than one transition probabilities from one state to another. Thereafter, a policy improvement methodology to analyze the asymptotic total expected energy consumption of the MS is taken into account.

A. DTMC Analysis with Rewards

In our model, a state of the process refers to the link in which the MS can operate, i.e., a WLAN, Ad hoc, or Cellular link. We denote these states by W, A, and C, respectively. There are certain transition probabilities between states obeying Markovian property, i.e., the conditional probability distribution of future states of the process depends only upon the present state, not on the sequence of events that preceded it. Hereafter, for the sake of generality of mathematical expression, we write states as W, A, and C; or 1, 2, and 3, respectively and these two sets of symbols will be used interchangeably.

A matrix $\mathbf{P}$, representing all the transition probabilities between states, can be written in elemental form as $[p_{ij}]$, $i, j = 1, 2, \ldots, N$. Here, i and j represent the row and column index, indicating the present and next state, respectively. As $\mathbf{P}$ is a Markovian matrix, we can write $\sum_j p_{ij} = 1$, $i = 1, 2, \ldots, N$.

When the system transfers (or does not transfer) from one state to another state, there is an associated reward/cost. The reward (or the cost) in our study is considered as the MS energy consumption per communicated byte in the corresponding link. Hence, we have a reward matrix, $\mathbf{R}$, for the system. As there is a reward value for each corresponding transition probability in $\mathbf{P}$, the size of $\mathbf{R}$ is the same as of $\mathbf{P}$. The elemental form for $\mathbf{R}$ can be written as $[r_{ij}]$, $i, j = 1, 2, \ldots, N$. For example, an element r_{12} of $\mathbf{R}$ implies the reward earned (or the cost paid) when the system moves from State-1 to State-2 with probability p_{12}.

Let q_i be the expected reward for the *next* transition if the current state is i. Then,

$$q_i = \sum_{j=1}^{N} p_{ij} r_{ij}, \; i = 1, 2, \ldots, N. \tag{1}$$

Furthermore, let $v_i(n)$ be the expected *total* reward in the next n transitions if the process starts in state i. Thus, we can write

$$\begin{aligned} v_i(n) &= \sum_{j=1}^{N} p_{ij}[r_{ij} + v_j(n-1)] \\ &= q_i + \sum_{j=1}^{N} p_{ij} v_j(n-1), \end{aligned} \tag{2}$$

where $i = 1, 2, \ldots, N$ and $n = 1, 2, \ldots$.

Eq. (2) recursively gives the total expected reward after n transitions. For example, $v_1(n)$ means the total expected reward after n transitions if the process starts in State-1, and $v_2(n)$ means the total expected reward after n transitions if the process starts in State-2; and so on.

It is important to note that in the above derived equations, there is just one possible direction of transition from a current state to a future state in the DTMC. This is distinct from an FMDP (to be elaborated below). Moreover, the equations progressively and cumulatively obtain the total expected reward for the process after n transitions. Therefore, $v_i(n)$ is an unbounded quantity. However, we are interested in a more valuable quantity, i.e., the *average reward per unit transition*, also referred to as gain. Therefore, an FMDP needs to be utilized when more than one possible directions of transition from each state exist, as described next.

B. FMDP Analysis with Rewards

An FMDP is a stochastic reinforcement learning technique [8] that satisfies the Markov property. It is defined by its state and action spaces and by one-step dynamic of the environment. The state space is composed of all the states that the process can have while the action space consists of all the possible actions that can be taken from these states. An action may (or may not) lead to a transition of the process from the current state to any of the other states *with certain probabilities*. For example, if an action is taken in State-2, the process may move to State-1 or State-3 with certain probabilities or may remain in State-2 with a probability. Thus, with a total transition probability of one, a certain action can lead the process to next/previous/current state. Therefore, given any state *and* action, there is a probability of transition. Thus, a state transition in our FMDP means the process moving with certain probability to another state based on the current state-action pair. Furthermore, given any current state and action, together with any next state, there is an expected value of the associated reward (due to the action taken). As the objective is to minimize this expected cost (which is basically the energy consumption of the MS in different links in this study), a policy needs to be optimized for this purpose. A policy means a set of all actions (one chosen action for each state) at a given instant. For the next iteration, a different (than previous) action may be chosen in a given state. Hence, a different policy. However, a randomly selected policy does not lead to minimal energy consumption. Therefore, we use policy iteration method to reach at the optimized policy.

Similar to the DTMC analysis with transition probability and reward matrices, the steady-state gain for an FMDP can be defined as

$$G = \sum_{i=1}^{N} \pi_i q_i, \; i = 1, 2, \ldots, N, \tag{3}$$

where π_i are the limiting values of the elements of probability transition matrix and can be easily obtained [9]. q_i can be obtained in a similar way as in Eq. (1).

Now, suppose we have more than one possible actions that can be taken in a state. Define $\mathbf{d} = d_i(n)$ as the policy[1] to use when the system is in state i. The task is to reach an optimum policy, i.e., the set of actions which leads to the least total expected reward (despite the initial policy). Thereafter, we improve the initially selected policy step by step. Therefore, policy iteration algorithm is used.

There are three steps in the policy iteration algorithm, i.e., value determination, policy improvement, and convergence test. These steps are described in the following.

B.1. Value Determination

In this step, our goal is to determine the values of v_i and G for each iteration.

We start from a given policy for a specific Markov chain with rewards, and obtain $v_i(n)$. For large n, taking into account the Bellman's Principle of Optamily[2], the asymptotic expression for $v_i(n)$ can be written as

$$v_i(n) = n\sum_{i=1}^{N} \pi_i q_i + v_i = nG + v_i,\ n = 1,2,\dots. \tag{4}$$

Now utilizing Eq. (4) in Eq. (2), we obtain

$$nG + v_i = q_i + \sum_{j=1}^{N} p_{ij}[(n-1)G + v_j],\ i = 1,\dots,N, \tag{5}$$

which leads to

$$G + v_i = q_i + \sum_{j=1}^{N} p_{ij} v_j,\ i = 1,\dots,N. \tag{6}$$

In the above expression, there are N linear simultaneous equations with $N+1$ unknown (i.e., v_i and G). The solution gives v_i and G at each iteration and these are called relative values of the policy because these values are used as the input for the next iteration in order to improve the policy.

B.2. Policy Improvement

This step is targeted at finding a better alternative/action for each state for the next iteration in order to improve the gain, G.

Let us start at stage n with a policy. We can then find the best alternative in the ith state at stage $n+1$ by using (see Eq. (6))

$$\arg\min_k\ q_i^k + \sum_{j=1}^{N} p_{ij}^k v_j(n),\ i = 1,\dots,N,$$

[1] $\mathbf{d}$ is a set of all possible actions that can be taken at time n in state i.

[2] A reader may refer to [10] and [11] for an in-depth study on Bellman's principle of optimality.

where k is simply an index for possible alternatives (i.e., actions).

For a large n, using Eq. (4) we can write the above expression as

$$\begin{aligned} & \arg\min_k \; q_i^k + \sum_{j=1}^{N} p_{ij}^k (nG + v_j) \\ = \; & \arg\min_k \; q_i^k + nG + \sum_{j=1}^{N} p_{ij}^k v_j, \; i = 1, \ldots, N. \end{aligned}$$

As nG is independent of the indices i, j, and k, we need to minimize only

$$\arg\min_k \; q_i^k + \sum_{j=1}^{N} p_{ij}^k v_j, \; i = 1, \ldots, N. \tag{7}$$

Now, the relative values, v_j, obtained in Eq. (6) can be applied in Expression (7) to obtain the values of G and v_i for policy improvement. Thus, we get new values of G and v_i (and correspondingly $\mathbf{d}$).

B.3. Convergence Test

Depending on the numerical values used for different parameters, the difference between the current and the previous values of v_i (i.e., v_i at transition n and v_i at transition $n-1$, for each i) may not converge to 0 with limited number of iterations. Therefore, a threshold on the difference between two consecutive values of v_i needs to be defined. If the difference falls below the threshold level, it indicates that the algorithm has converged; and we stop with obtained values of v_i. Otherwise, next iteration is performed.

C. Actions and Station Transitions in an MS using FMDP

Based on the necessary foundation developed in Eqs. (1) through (7), we can associate the analysis to our scenario for the MS. Fig. D.2 illustrates the FMDP used for the MS with associated actions and state transitions. To avoid pictorial complexity, the possible state transitions are shown separately for each state. There are three possible states that the MS can have (i.e., $1 \equiv W$, $2 \equiv A$, and $3 \equiv C$), shown as bigger circles. Furthermore, there are three possible actions, shown as smaller circles, that can be chosen in each state[3]. The actions in our study are defined as follows:

- Action-a: The MS operates in the High Energy Saving (HES) mode, where data transfer speed and delay have lowest priority,

[3]For illustration simplicity, the action and the state spaces in terms of transition options have been deliberately kept small. Larger spaces will lead to more complex policies.

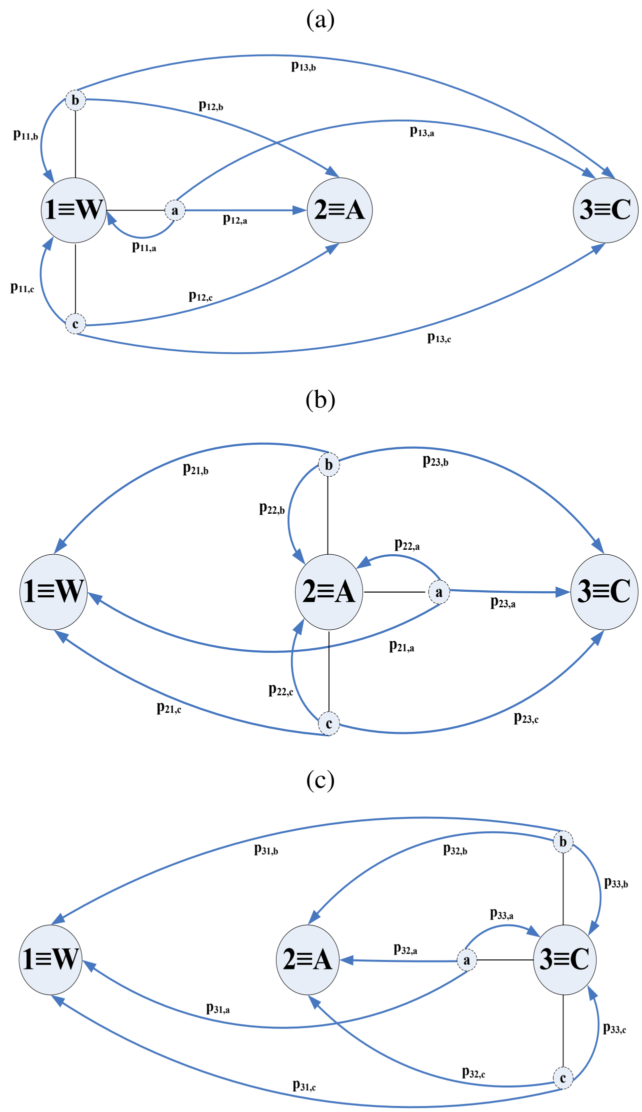


Figure D.2: Partial FMDPs for the MS: initial state at (a) W, (b) A, and (c) C.

- Action-b: The MS operates in the Low Energy Saving (LES) mode, where data transfer speed and delay have highest priority, and

- Action-c: The MS operates in the Energy Ignorant (EI) mode, where it decides to join a link based on the available data rates. The MS does not care about which link consumes more or less energy.

As mentioned earlier, an action may (or may not) lead the system to another state with certain probabilities. For example $p_{WC,b}$ means the probability that the MS changes the link from W to C if Action-b is selected in State-W. Similarly, $p_{WW,a}$ implies the probability that the MS keeps on operating in the link W if Action-a is chosen in State-W; and so on.

In order to make the above mentioned terminology consistent with our generalized analysis in Eqs. (1) through (7), we write $p_{WC,b}$ as $p_{13,b}$ and $p_{WW,a}$ as $p_{11,a}$; and so on. Moreover, N is equal to 3 for the system under consideration. It is also worth noting that $\sum_{j=1}^{N} p_{ij,X} = 1$, where $i = 1, \ldots N$ and $X \in \{a, b, c\}$.

The rewards associated with each state-action pair can also be represented in a similar fashion as mentioned in the above paragraphs. For example, $r_{12,a}$ implies the reward when the system moves from link W to link A when Action-a is selected; and so on.

In summary, an MS may start from any connection, depending on its location, data rate requirement or remaining battery level. It then can interchangeably operate in any of the available links and take any of the possible actions based on the status of its communication progress and the values of these parameters. For example, an MS starts initially its communication in State-W and gives higher probability to Action-c. As it moves away from the AP or the battery level goes down below a certain threshold, it prefers to take Action-a in order save energy while still staying as connected. The transition probabilities may in different cases, be a function of a parameter between the involved transceivers as well as the amount of possible data transfer rate for the given links. Therefore, the MS may switch links with certain probabilities by taking the mentioned actions in order to improve the overall long-term energy consumption through policy iteration, as described in Sec. III.B.

IV. Numerical Results and Discussions

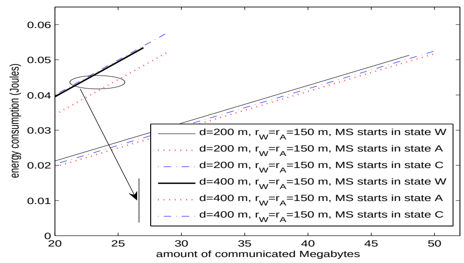
This section presents the numerical results achieved from the analysis of the scheme. A custom-built simulation tool is developed in MATLAB to numerically evaluate the system performance with different parameter configurations summarized in Table D.1 [2], [12], and [13]. The basic rate in the table refers to the rate at which the control frames are communicated.

Table D.1: Parameter configuration.

Parameter	Value	Parameter	Value
basic rate	6 Mbps	T_{CTS}	$14*8/$basic rate
$R_a = R_w$	54 Mbps	T_{SIFS}	$10\,\mu s$
data frame size	300 bytes	T_{DIFS}	$50\,\mu s$
P^E	0.635×10^{-7} W	T_{RTS}	$20*8/$basic rate
$E[T_{bo}]$	$80\,\mu s$	T_{beacon}	$179*8/$basic rate
R_c	2 Mbps	T_{ACK}	$14*8/$data rate

Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) has been used as the underline MAC protocol for energy consumption calculations in the HDR

(a)



(b)

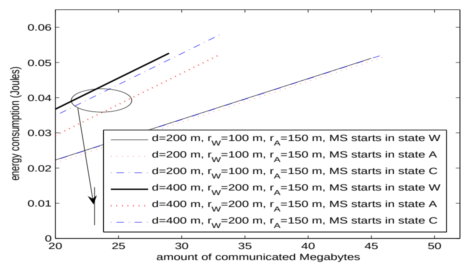


Figure D.3: Energy consumption using DTMC when the MS starts in state W, A, or C, for various distances between the MS and the BS, AP, or RS.

connections. For the WLAN link, the energy consumed by the MS is calculated by considering the MS as operating in the CFP only. PCF decides about the duration of CP and CFP as well as the association or disassociation of the MS with the WLAN link. Data rates for the cellular, the ad hoc, and the WLAN link are respectively denoted by R_c, R_a, and R_w. Corresponding distances between the MS and the BS, the AP, and the RS are respectively denoted by d[4], r_W, and r_A. The representative values for d are 200, 400, 500, and 600 meters. The values of r_W are 100, 150, and 200 meters, respectively; while for r_A they are taken as 50, 100, and 150 meters. The values of path-loss coefficient are taken as 2.5 for the cellular link; while for the ad hoc and the WLAN link, they are taken as 4. The expected back-off time for the CP is taken as 80 microseconds.

In all the figures from Fig. D.3 through Fig. D.5, the MS energy consumption is plotted versus the amount of communicated bytes.

In the following, we first investigate the numerical results obtained for the MS energy consumption using DTMC. Later on, FMDP-based energy consumption is discussed. Lastly, discussions on policy convergence are also presented.

A. MS Energy Consumption using DTMC with Rewards

Figs. D.3(a) and D.3(b) illustrate the MS battery energy consumption (and implicitly, the MS battery lifetime) for given starting states. The x-axes represent the

[4] In Figs. D.3, D.4, and D.5, d (without subscript) denotes the cellular link distance between the MS and the BS. This is different from $\mathbf{d} = d_i(n)$ (as mentioned in Sec. III.B) which refers to a policy.

amount of communicated bytes and the y-axes represent the consumed MS battery energy.

In Fig. D.3(a), by comparing the upper set of plots (indicated by the ellipse) with the lower set of plots, we observe that, for the same amount of energy consumption, the MS can communicate relatively smaller amount of bytes when the cellular link distance is increased from 200 meters to 400 meters, while keeping the ad hoc and WLAN link distances fixed. This is because that the MS needs to transmit at a higher power level to reach the BS, which increases the overall energy consumption, and thus the battery depletes quicker. As the probabilities of transition in the DTMC are fixed, the average fraction of time the MS remains in State-C is the same for both $d = 200$ meters and $d = 400$ meters. However, $d = 400$ meters causes a larger amount of energy consumption. Furthermore, as the MS switches between the states, the lifetime of the MS battery is brought down due to higher cellular link distance no matter what state the system starts with. This is because of a higher cost in State-C with a higher value of d. Hence, the upper set of plots in Fig. D.3(a).

Comparing the upper set of plots with the lower one within Fig. D.3(a), an interesting observation is that the MS battery consumes lower amount of energy when the process starts in State-W (see the solid line in the plots) compared with when the process starts in State-A. However, the difference between energy consumption when the process starts in State-W (or State-C) and when the process starts in State-A is significant. This is due to the reason that a larger value of link distance in State-C as well as the fixed state-transition probability contribute more to the overall energy consumption. Hence, the long-term consumption starting in State-W approaches the long-term consumption starting in State-C, although the WLAN link consumes relatively more energy for the same values of d and r_W. Moreover, as the transition probabilities are fixed, the MS has to change state (i.e., link) and thus it cannot avoid switching to a high energy-consuming link.

Furthermore, comparing the upper set of plots with the lower set within Fig. D.3(a), we also observe that, for a given amount of consumed energy, when the cellular link distance is decreased the MS can communicate more data when it starts from State-C (solid plot in the lower set of curves is above the dash-dot plot) as compared with when it starts from State-W. The reason for this effect is similar as described in the above two paragraphs.

Moreover, note in the upper set of plots in Fig. D.3(b) that the value of r_W is 200 meters while in the lower set it is 100 meters. Furthermore, d is 200 and 400 meters in the upper and lower set of plots, respectively. As obvious, the MS battery depletes quicker if it starts in State-W, because of the initial probability assignment

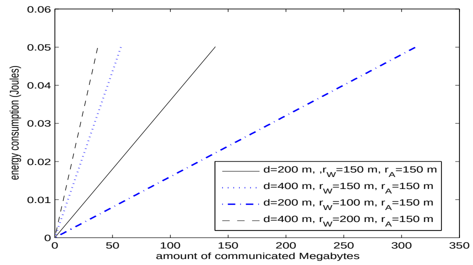


Figure D.4: Minimized energy consumption using FMDP for various distances between the MS and the BS, AP, or RS.

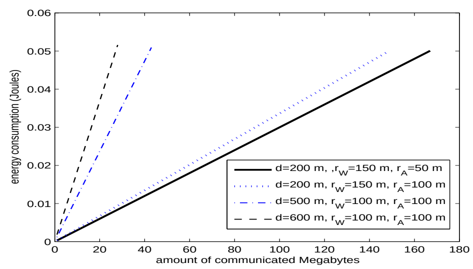


Figure D.5: Minimized energy consumption using FMDP for various distances between the MS and the BS, AP, or RS.

of the MS being in State-W as well as the higher values of d and r_W. Furthermore, in the lower set of plots, there is a negligible difference between the plots for the MS starting in states W and C. This is because that with even lower value of r_W ($= 100$ meters), the energy consumption is relatively lower and thus the plot for State-W approaches the plot for State-C (when d is also smaller).

From Fig. D.3 (both (a) and (b)), an overall finding is that the DTMC-based approach enables us to estimate the battery lifetime for an MS operating in an HN. However, the picture of the MS energy consumption can be different for different *starting states* of the system.

A drawback of the the DTMC-based approach is that the system has fixed transition directions. Therefore, the feasibility of multiple transition directions towards minimized energy consumption is also explored, as discussed below.

B. MS Energy Consumption using FMDP with Rewards

Figs. D.4 and D.5 illustrate the MS energy consumption versus the amount of communicated data using FMDP-based approach. It is pertinent to mention that as the MS energy consumption is jointly minimized and averaged for long-term for all the states, unlike Fig. D.3, Figs. D.4 and D.5 do not separately illustrate the MS energy consumption when starting in a certain state.

As illustrated in Fig. D.4, for smaller values of d and r_W (while keeping r_A as constant), and for the same amount of battery energy consumption, the MS can communicate a significantly higher amount of bytes. The reason for this effect is twofolds. First, due to the decision making capability of the FMDP and the flexibility of available options, the MS chooses the actions/alternatives which further reduce (i.e., minimize) the long-term energy consumption. Second, because of the shorter values of cellular and WLAN link distances, the battery lifetime becomes longer. Furthermore, by comparing the solid plot and the bold dash-dot plot in Fig. D.4, we observe that increasing the WLAN link distance from 100 meters to 150 meters, while keeping $d = 200$ meters, significantly decreases the amount of communicated bytes. This is because that the WLAN connection consumes more energy per meter increase of distance between transceivers compared with the cellular connection for similar values of path-loss coefficient. Therefore, the energy consumption of State-W overwhelms the other states' energy consumption.

Furthermore, Fig. D.5 also illustrates the effects of distance between transceivers on the minimized energy consumption of the MS in the HN. In this figure, as expected, smaller values of d and r_A increase the optimized overall amount of communicated bytes. The reasons for this effect are the shorter connection distance as well as the overall minimization of MS energy consumption due to the selection of least energy consuming alternatives in the FMDP, as described in the above paragraph also. This means that by using the FMDP-based approach, the MS reaches the optimum policy for minimum energy consumption.

Moreover, though not explicitly shown in Fig. D.3 due to scale, the initial phase of the curves is not linear. That is, the expected reward starting from each state behaves nonlinearly. However, after only a few transitions, the difference between expected rewards (for any two given starting states) becomes constant. Hence, the linear part of the plots. On the other hand, in Figs. D.4 and D.5, as the MS energy is jointly minimized and averaged (for all starting states), there is just one plot for each given set of distance parameters. Therefore, the effects of starting states are implicitly absorbed within a plot in Figs. D.4 and D.5.

Comparing Fig. D.3 with Figs. D.4 and D.5, there is a clear difference between the amount of bytes communicated by using the DTMC-based and FMDP-based approaches. Using the FMDP-based approach, the MS battery can last longer. This is because that in the FMDP-based approach with multiple available options, the MS chooses actions such that the long-term energy consumption is minimized. That is, the MS more often operates in (or return to) the HES mode (by choosing Action-a) and thus save energy. In other words, due to the availability of flexible options, the

MS can select optimum actions through progressive policy improvement, as already described in Sec. III.B.

C. Convergence of Policy

As in the FMDP-based approach, the process starts with an initial policy and is converged to an optimal policy through iterations, the convergence time of the employed policy iteration algorithm is also of interest. The convergence of the policy iteration algorithm is mainly dependent on the size of the system, i.e., on both the number of states and on the possible actions that can be chosen. For example, for a system with l states and m actions, the total number of possible policies becomes $l \times (m-1)!$. This means that with a modest increase in the number of states (or actions), the total number of possible policies can increase very quickly. However, in our model, as there are only 3 states and 3 possible actions, the algorithm is observed to converge quickly (within only $8 \sim 11$ iterations).

V. Conclusions and Future Work

In this paper, we have proposed two approaches to estimate the long-term energy consumption of MS battery in a heterogeneous network comprising of WLAN, ad hoc, and cellular networks. By using a DTMC-based approach, the MS's battery lifetime can be predicted given that it starts from a certain link. By using the policy improvement methodology in an FMDP-based approach, the MS has more flexibility in choosing a low energy-consuming network over long-term. This approach minimizes the average MS energy consumption (over all network connections); and thus ensures a longer battery lifetime as compared with the DTMC-based approach.

As our future work, the proposed approaches will be implemented and tested for battery energy consumption, based on real-life network. Moreover, the effects of the approaches will also be observed on the RS battery lifetime in the involved ad hoc network. Joint optimization of data-transfer speed and MS energy consumption for heterogeneous traffic will also be explored.

References

[1] N. Balasubramanian, A. Balasubramanian, and A. Venkataramani, "Energy Consumption in Mobile Phones: A Measurement Study and Implications for Network Applications," in *Proc. ACM/USENIX Internet Measurement Conference (IMC)*, Chicago, USA, November 2009.

[2] Z. H. Abbas and F. Y. Li, "Power Consumption Analysis for Mobile Stations in Hybrid Relay-assisted Wireless Networks," in *Proc. 5th IEEE International*

Symposium on Wireless Pervasive Computing (ISWPC), Modena, Italy, May 2010.

[3] Y. W. Chung, "Investigation of Energy Consumption of Mobile Stations for Instant Messaging Services," in *Proc. 10th International Symposium on Autonomous Decentralized Systems (ISADS)*, Kobe, Japan, June-July 2011.

[4] T. L. Cheung, K. Okamoto, F. Maker, X. Li, and V. Ajella, "Markov Decision Process (MDP) Framework for Optimizing Software on Mobile Phones," in *Proc. International Conference on Embedded Software (EMSOFT)*, Grenoble, France, October 2009.

[5] IEEE Std. IEEE 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE-SA, June 2007. doi:10.1109/IEEESTD.2007.373646.

[6] M. S. Gast, *802.11 Wireless Networks The Definitive Guide*, 2nd edition, O'Rielly and Associates Inc., 2005.

[7] Z. H. Abbas and F. Y. Li, "Distance-related Energy Consumption Analysis for Mobile/Relay Stations in Heterogeneous Wireless Networks," in *Proc. 7th IEEE International Symposium on Wireless Communication Systems (ISWCS)*, York, UK, September 2010.

[8] F. S. Hillier and G. J. Lieberman, *Introduction to Operations Research*, 8th edition, McGraw-Hill Companies Inc., New York, USA, 2005.

[9] F. Gebali, *Analysis of Computer and Communication Networks*, Springer Science+Business Media, New York, USA, 2008.

[10] R Bellman, "The Theory of Dynamic Programming," *Bulliten of the American Mathematical Society*, Vol. 60, 1954.

[11] M. Sniedovich, "A New Look at Bellman's Principle of Optimality," *Journal of Optimization Theory and Applications*, Vol. 49, No. 1, April 1986.

[12] L. M. Feeney and M. Nilsson, "Investigating the Energy Consumption of a Wireless Network Interface in an Ad Hoc Networking Environment," in *Proc. IEEE INFOCOM*, Anchorage, AK, USA, April 2001.

[13] C. Levy-Bencheton and G. Villemaud, "Power Consumption Optimization in Multi-mode Mobile Relay," in *Proc. 12th European Conference on Wireless Technology (EuWiT)*, Rome, Italy, September 2009.